UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1272

Mass Spectrometry and Affinity Based Methods for Analysis of Proteins and Proteomes

MÅRTEN SUNDBERG

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2015

ISSN 1651-6214
ISBN 978-91-554-9300-4
urn:nbn:se:uu:diva-259623

Dissertation presented at Uppsala University to be publicly examined in C4:305, BMC, Husargatan 3, Uppsala, Friday, 25 September 2015 at 10:15 for the degree of Doctor of Philosophy. The examination will be conducted in Swedish. Faculty examiner: Professor Christer Wingren (Lund University).

Abstract
Sundberg, M. 2015. Mass Spectrometry and Affinity Based Methods for Analysis of Proteins and Proteomes. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 1272. 82 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-554-9300-4.

Proteomics is a fast growing field and there has been a tremendous increase of knowledge the last two decades. Mass spectrometry is the most used method for analysis of complex protein samples. It can be used both in large scale discovery studies as well as in targeted quantitative studies. In parallel with the fast improvements of mass spectrometry-based proteomics there has been a fast growth of affinity-based methods. A common challenge is the large dynamic range of protein concentrations in biological samples. No method can today cover the whole dynamic range. If affinity and mass spectrometry-based proteomics could be used in better combination, this would be partly solved. The challenge for affinity-based proteomics is the poor specificity that has been seen for many of the commercially available antibodies. In mass spectrometry, the challenges are sensitivity and sample throughput. In this thesis, large scale approaches for validation of antibodies and other binders are presented. Protein microarrays were used in four validation studies and one was based on mass spectrometry. It is shown that protein microarrays can be valuable tools to check the specificity of antibodies produced in a large scale production. Mass spectrometry was shown to give similar results as Western blot and Immunohistochemistry regarding specificity, but did also provide useful information about which other proteins that were bound to the antibody.

Mass spectrometry has many applications and in this thesis two methods contributing with new knowledge in animal proteomics are presented. A combination of high affinity depletion, SDS PAGE and mass spectrometry revealed 983 proteins in dog cerebrospinal fluid, of which 801 were marked as uncharacterized in UniProt. A targeted quantitative study of cat serum based on parallel reaction monitoring showed that mass spectrometry can be an applicable method instead of ELISA in animal proteomic studies. Mass spectrometry is a generic method and has the advantage of shorter and less expensive development costs for specific assays that are not hampered by cross-reactivity.

Mass spectrometry supported by affinity based applications will be an attractive tool for further improvements in the proteomic field.

Keywords: Mass spectrometry, proteomics, microarray, protein, antibody, antigen, affinity, validation

Mårten Sundberg, Department of Chemistry - BMC, Analytical Chemistry, Box 599, Uppsala University, SE-75124 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-554-9300-4
urn:nbn:se:uu:diva-259623 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-259623)

Till min fantastiska familj

Katarina, Viggo och Joel

Main Supervisor:	**Margareta Ramström**, Associate Professor Department of Chemistry - BMC, Analytical Chemistry Uppsala University Uppsala, Sweden
Assistant supervisor:	**Jonas Bergquist**, Professor Department of Chemistry - BMC, Analytical Chemistry Uppsala University Uppsala, Sweden
Faculty opponent:	**Christer Wingren**, Professor Department of Immunotechnology Lund University Lund, Sweden
Examining committee:	**Amelie Eriksson Karlström**, Professor Division of Protein Technology, School of biotechnology Royal Institute of Technology Stockholm, Sweden
	Peter Lindblad, Professor Department of Chemistry - Ångström Laboratory, Molecular Biomimetics; Microbial Chemistry Uppsala University Uppsala, Sweden
	Michael Lund Nielsen, Professor Novo Nordisk Foundation Center for Protein Research University of Copenhagen Copenhagen, Denmark
	Corine Sandström, Professor Department of Chemistry and Biotechnology Swedish University of Agricultural Sciences Uppsala, Sweden
	Per Matsson, Associate Professor Thermo Fisher Scientific Uppsala, Sweden

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I. Nilsson, P., Paavilainen, L., Larsson, K., Odling, J., **Sundberg, M**., Andersson, A. C., Kampf, C., Persson, A., Al-Khalili Szigyarto, C., Ottosson, J., Bjorling, E., Hober, S., Wernerus, H., Wester, K., Ponten, F., and Uhlen, M. (2005) Towards a human proteome atlas: high-throughput generation of mono-specific antibodies for tissue profiling, *Proteomics 5(17)*, 4327-4337.

II. Eriksson, C., Agaton, C., Kange, R., **Sundberg, M**., Nilsson, P., Ek, B., Uhlen, M., Gustafsson, M., and Hober, S. (2006) Microfluidic analysis of antibody specificity in a compact disk format, *J Proteome Res 5(7)*, 1568-1574.

III. Schwenk, J. M., Lindberg, J., **Sundberg, M**., Uhlen, M., and Nilsson, P. (2007) Determination of binding specificities in highly multiplexed bead-based assays for antibody proteomics, *Mol Cell Proteomics 6(1)*, 125-132.

IV. Sjoberg, R*., **Sundberg, M***., Gundberg, A., Sivertsson, A., Schwenk, J, M., Uhlen, M., and Nilsson, P (2012) Validation of various affinity reagents using protein microarrays. *N Biotechnol. 29(5),* 555-63.

V. Andersson, S., **Sundberg, M**., Pristovsek, N., Wang, J., Jonsson, P., Clausson, C-M., Zieba, A., Ramström, M., Söderberg, O., Williams, C., and Asplund, A. Estrogen receptor beta profiling in human tissues following extensive antibody validation. Manuscript.

VI. **Sundberg, M**., Strage, E. M., Bergquist, J., Ström Holst, B., and Ramström, M. Quantitative and selective analysis of feline growth related proteins using parallel reaction monitoring high resolution mass spectrometry. Manuscript.

VII. **Sundberg, M**., Bergquist, J., and Ramström, M. (2015) High-abundant protein depletion strategies applied on dog cerebrospinal fluid and evaluated by high-resolution mass spectrometry. *Biochemistry and Biophysics Reports 3,* 68-75.

Reprints were made with permission from the respective publishers.
*These authors contributed equally to this work.

Author Contribution

Paper I: Took part in the planning, did some of the experiments and took part of writing the paper.

Paper II: Took part in the planning, did some of the experiments and took part of writing the paper.

Paper III: Took part in the planning, did some of the experiments and took part of writing the paper.

Paper IV: Took part in the planning, did the major part of the experiments and was one of the main writers of the paper.

Paper V: Took part in the planning, did some of the experiments and wrote part of the paper.

Paper VI: Did the planning and performed the experiments of all mass spectrometry-related experiments. Wrote the paper.

Paper VII: Did the planning and performed the experiments. Wrote the paper.

Related Publications not Included in the Thesis

Colwill, K., Persson, H., Jarvik, N. E., Wyrzucki, A., Wojcik, J., Koide, A., Kossiakoff, A. A., Koide, S., Sidhu, S., Dyson, M. R., Pershad, K., Pavlovic, J. D., Karatt-Vellatt, A., Schofield, D. J., Kay, B. K., McCafferty, J., Mersmann, M., Meier, D., Mersmann, J., Helmsing, S., Hust, M., Dubel, S., Berkowicz, S., Freemantle, A., Spiegel, M., Sawyer, A., Layton, D., Nice, E., Dai, A., Rocks, O., Williton, K., Fellouse, F. A., Hersi, K., Pawson, T., Nilsson, P., **Sundberg, M**., Sjoberg, R., Sivertsson, A., Schwenk, J. M., Takanen, J. O., Hober, S., Uhlen, M., Dahlgren, L. G., Flores, A., Johansson, I., Weigelt, J., Crombet, L., Loppnau, P., Kozieradzki, I., Cossar, D., Arrowsmith, C. H., Edwards, A. M., and Graslund, S. (2011) A roadmap to generate renewable protein binders to the human proteome, *Nat Methods 8(7)*, 551- 558.

Rimini, R., Schwenk, J. M., **Sundberg, M**., Sjoberg, R., Klevebring, D., Gry, M., Uhlen, M., and Nilsson, P. (2009) Validation of serum protein profiles by a dual antibody array approach, *J Proteomics 73(2)*, 252-266.

Lundberg, E., **Sundberg, M**., Graslund, T., Uhlen, M., and Svahn, H. A. (2007) A novel method for reproducible fluorescent labeling of small amounts of antibodies on solid phase, *J Immunol Methods 322(1-2)*, 40-49.

Uhlen, M., Bjorling, E., Agaton, C., Szigyarto, C. A., Amini, B., Andersen, E., Andersson, A. C., Angelidou, P., Asplund, A., Asplund, C., Berglund, L., Bergstrom, K., Brumer, H., Cerjan, D., Ekstrom, M., Elobeid, A., Eriksson, C., Fagerberg, L., Falk, R., Fall, J., Forsberg, M., Bjorklund, M. G., Gumbel, K., Halimi, A., Hallin, I., Hamsten, C., Hansson, M., Hedhammar, M., Hercules, G., Kampf, C., Larsson, K., Lindskog, M., Lodewyckx, W., Lund, J., Lundeberg, J., Magnusson, K., Malm, E., Nilsson, P., Odling, J., Oksvold, P., Olsson, I., Oster, E., Ottosson, J., Paavilainen, L., Persson, A., Rimini, R., Rockberg, J., Runeson, M., Sivertsson, A., Skollermo, A., Steen, J., Stenvall, M., Sterky, F., Stromberg, S., **Sundberg, M**., Tegel, H., Tourle, S., Wahlund, E., Walden, A., Wan, J., Wernerus, H., Westberg, J., Wester, K., Wrethagen, U., Xu, L. L., Hober, S., and Ponten, F. (2005) A human protein atlas for normal and cancer tissues based on antibody proteomics, *Mol Cell Proteomics 4(12)*, 1920-1932.

Janzi, M., Odling, J., Pan-Hammarstrom, Q., **Sundberg, M**., Lundeberg, J., Uhlen, M., Hammarstrom, L., and Nilsson, P. (2005) Serum microarrays for large scale screening of protein levels, *Mol Cell Proteomics 4(12)*, 1942-1947.

Contents

Abbreviations

ABP	Albumin binding protein
ACS	Acute coronary syndrome
AGC	Automatic gain control
AQUA	Absolute quantification
BSA	Bovine serum albumin
CDR	Complementarity determining regions
CID	Collision induced dissociation
CIMS	Context-independent motif specific
CRM	Charge residue model
CSF	Cerebrospinal fluid
C18	Octadecyl carbon chain
DARPin	Designed ankyrin repeat protein
DDA	Data-dependent acquisition
DIA	Data-independent acquisition
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ELISA	Enzyme-Linked Immunosorbent assay
ESI	Electrospray ionization
ETD	Electron transfer dissociation
Fab	Fragment, antigen binding
Fc	Fragment, crystallizable
FlexiQuant	Full-length expressed stable isotope-labeled proteins for quantification
FTICR	Fourier transform ion cyclotron resonance
Fv	Fragment, variable
GPS	Global proteome survey
HCD	Higher-energy collisional dissociation
HILIC	Hydrophilic-interaction chromatography
His_6	Hexahistidine
HPA	Human protein atlas
HPLC	High-performance liquid chromatography
HUPO	Human proteome organization
IAA	Iodoacetamide
ICAT	Isotope coded affinity tag
IEM	Ion evaporation model
IF	mmunofluorescence
Ig	Immunoglobulin

IHC	Immunohistochemistry
IEX	Ion exchange
ISET	Integrated selective enrichment target
iTRAQ	Isobaric tags for relative and absolute quantification
LC	Liquid chromatography
LTQ	Linear trap quadrupole
MALDI	Matrix-assisted laser desorption ionization
MRM	Multiple reaction monitoring
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MSIA	Mass spectrometric immunoassay
MS/MS	Tandem mass spectrometry
m/z	Mass over charge
NHS	*N*-Hydroxysuccinimide
NCE	Normalized collision energy
PCR	Polymerase chain reaction
PrEST	Protein epitope signature tag
PRM	Parallel reaction monitoring
PSAQ	Protein standard absolute quantification
PSM	Peptide spectrum match
PTM	Post-translational modification
PVDF	Polyvinylidene fluoride
Q	Quadrupole
QconCAT	Quantification concatamer
QPrEST	Quantification protein epitope signature tag
RIA	Radioimmunoassay
RISQ	Recombinant isotope labeled and selenium quantified
RNA	Ribonucleic acid
RP-HPLC	Reversed phase-high-performance liquid chromatography
RT-PCR	Reverse transcriptase- Polymerase chain reaction
scFv	Single-chain variable fragment
SCX-HPLC	Strong cation exchange- high-performance liquid chromatography
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGC	Structural genomics consortium
SILAC	Stable isotope labeling of amino acids in cell culture
SISCAPA	Stable isotope standards and capture by anti-peptide antibodies
SRM	Selected reaction monitoring
TMT	Tandem mass tag
TOF	Time-of-flight
TXP	Triple X Proteomics
UPLC	Ultra-performance liquid chromatography

1. Introduction

The start of life on earth goes billions of years back in time. Still it was not until very recently that humans began to explore and slowly achieve a better understanding of the complexity of the building blocks of life. The human body is composed of a very large number of different cells. The genetic code is the basis of all living cells and the information in the code controls all activities of the living cells. The total length of the human genome is over 3 billion base pairs and the nucleotides are folded into a double helix spiral, the deoxyribonucleic acid, DNA [1]. DNA is the template for proteins, which are essential parts of all living organisms involved in most of their biological processes. Proteins cannot directly be produced from the DNA strand, there is another molecule needed called mRNA. This is an intermediate step from the DNA code to protein. This is called the central dogma [2]. It involves a transcription from DNA to pre-mRNA, which can be modified by different splicing of the sequence to get mRNA. Finally there is a translation from mRNA to protein. Today it is anticipated that the human proteome is based on about 20.300 protein coding genes [3]. That number is not the final answer to how many proteins the human body consists of. Proteins can have different isoforms due to mRNA splicing. However, it has been shown that one coding gene, one transcript is the most common [4]. Added to this, there is a large number of different post-translational modifications that makes the number of different functional proteins much higher [5]. Many of these are yet of unknown function and some of them might be key targets in the medicinal products that will be produced tomorrow. One of the challenges is to characterize these proteins to generate more knowledge about their involvement both in the processes in a healthy human body and in one that suffers from a disease.

The concept proteome was defined by Marc Wilkins during his doctoral studies in the 1990s [6] as a way to classify the study of the proteins of humans and other species. Proteomics is thus the large scale research concerning proteins and includes a large number of different techniques that can be used to find proteins in different samples. The exploration of proteins and the systems they are integrated in is of outmost importance in many research disciplines such as medicine, biochemistry and biology. One challenge in the proteomic research field is the large dynamic range of protein concentration that is found in e.g. the plasma proteome, which is up to 12 orders of magnitude [7]. In close relation to that is the sample handling and sample prepara-

tion. There is a need for establishing sensitive techniques that can detect low abundant proteins. Today there are mainly two different ways to approach this, mass spectrometry (MS) [8, 9] and affinity based methods. Regarding affinity based methods, there must be affinity reagents available that are specific. There are initiatives to make it easier for researchers to provide as much validation data as possible regarding antibodies and antigens. Two such internet based portals are antibodypedia.org [10] and pabmabs.com [11]. Mass spectrometry, on the other hand, can be performed as an undirected search for proteins in a sample. Since 2001, the human genome is sequenced [12, 13] and all information is now accessible through web-based databases such as UniProt and Ensembl [3, 14]. The number of completely sequenced species is constantly growing, benefiting both affinity and mass spectrometry proteomics. Even if the genome sequence does not reveal everything about the different proteomes it is a very good starting point for large scale proteomics studies. Without the enormous improvements in DNA sequencing, the proteomic field would not have developed as fast as it has during the last decade. Even if the development has been impressive, the field of proteomics, is still hampered by lack of sufficient sensitivity and specificity to uncover a whole proteome by a single method.

At the Human Proteome Organization (HUPO) conference in Madrid 2014, the keynote speaker Mathias Uhlén summarized his thoughts of what is most exciting in the proteomics field today. He formulated this as the "wonderful marriage between affinity and mass spectrometry based proteomics". This statement is a very good description of the aim of this thesis. It is of outmost importance that different fields of proteomics can co-operate to clarify the complexity of not only the human body, but also other species. This does not mean that the range of methods in the proteomics field must be reduced, but there has to be better ways of combining them. This will result in new analytical tools capable of investigating complex biological systems.

There have been tremendous improvements and discoveries in the proteomics research field the last decade. The first drafts of the human proteome, based on mass spectrometry [15, 16] and affinity [17] were released in May, 2014 and January, 2015, respectively. None of them claims to present the complete story of the human proteome and each study provides complementary and overlapping information about the Human Proteome. All information is publically available, but the challenge is to present it so that more information can be extracted from this fantastic resource. As has been seen with the human genome, we will see a continuous upgrading of the human proteome. Compared to the genomics field, there is still a lot that has to be done in the proteomics world e.g. improvements of analysis tools and infrastructure for data handling, storage and sharing of data [18]. Another important challenge in the proteomics field is to find and categorize different modifications that can be done by purpose or by chance (environmental effects, diseases, gender, etc) [19]. One such specific area is post-translational

modifications (PTM). Several hundred different PTMs have been reported and a gigantic work is ongoing to get an understanding of their biological functions [20-23]. If we together can find out the function of all different proteins, there is a fantastic opportunity for inventions of diagnostic tools that can classify differences that can appear in proteins at different disease stages. Hopefully resulting in earlier diagnoses as a result and consequently give a better chance to prevent different diseases like e.g. cancer, Alzheimer's and many more. There is a lot of work required to get there, but we proceed in the right direction with an increasing speed.

This thesis focuses on method developments of research tools within proteomics. One major problem in the affinity proteomics field is all the time and money spent on antibodies or other kinds of affinity binders that do not live up to the expectations [24-27]. Better validation strategies are needed and **Paper I-V** are contributions to this. Another issue within the proteomics field is the need of more specific and sensitive methods that could be used in multi species. Mass spectrometry is much easier to adapt to multi species which can save costs. The risk of cross-reactivity in mass spectrometry is limited compared to affinity-based proteomics and it should give a lot of added value to use more mass spectrometry in animal proteomics. To deal with this challenge there has to be a more multidisciplinary workflow than used today. Mass spectrometry and affinity proteomics must learn from each other and complement each other's shortcomings. Contributions to that are presented in **Paper VI and VII**.

2. Affinity Proteomics

A Diverse Field – History and Short Summary of Methods

Today there is a large number of different assays based on affinity reagents such as antibodies. A today widely used method was introduced in 1941 by Albert Coons and his coworkers [28]. The method was Immunohistochemistry which is a visualization method using antibodies *in situ* in tissue sections. A method that has become known as the first immunoassay was presented in 1959 by Rosalyn Yalow and Solomon Berson [29]. The method was given the name radioimmunoassay (RIA). Next revolution in the area was the introduction of enzyme-linked immunosorbent assay (ELISA) Peter Perlmann and Eva Engvall 1971 [30]. 1975 the concept monoclonal antibody was launched [31] and after that there have been many more publications presenting different kinds of binders and many of them will be presented later. Another achievement was the introduction of Western blot (WB) 1981[32]. This is just a few but very important discoveries among all that have been launched the last decades. There has been an incredible development and the release of the human genome in 2001 opened up new possibilities with in the proteomic field. Large-scale protein mapping methods such as protein arrays and mass spectrometry have the last decade increased protein knowledge in a magnificent way. Today, there is an amazing palette of methods to identify and characterize proteins. If used wisely, we would be able to make new discoveries that may reveal the mechanism behind many of today's worst diseases. This can then hopefully be used for new diagnostic methods and medications. To get there, new collaborations and strategies have to be implemented. Much has been performed in the past two decades, but unfortunately, the new insights have not resulted in as many new diagnostics or treatments that could have been expected.

Affinity Reagents Used in Protein Microarrays

Many types of molecules can be used for protein microarrays such as antibodies, recombinant protein, antibody fragment and different kinds of scaffolds [33]. In what context the protein microarray is going to be used determines what kind of capture reagents to use. In this thesis the focus is put on the antigen protein microarray as a tool for validation of produced antibodies or other affinity reagents (**Paper I-IV**). In antigen microarrays used in this thesis, the capture reagent is the antigen that the affinity reagent is produced against. If the array is inverted different kinds of affinity binders is printed and used as capturing agents, such as antibody arrays used for protein detection in different kinds of body fluids.

Antigens

An antigen is usually described as a substance foreign to the host organism. The substance will trigger an immune response that produces antibodies targeting that substance. The region of antibody recognition of an antigen is referred to as an epitope. Epitopes can be linear or conformational and it is important that the antibody can attain the epitope to be able to bind. Changes of epitope conformation can occur due to environmental change, e.g. denaturation. The ability of an affinity reagent to recognize the epitope in a native or denatured state can either be an advantage or a disadvantage.

The most common antigens used for production of antibodies are polysaccharides, lipids, nucleic acids and the most used, proteins. When antibodies are produced in host animals there are usually three different types of antigens used: recombinant full-length proteins, other recombinant proteins and small peptides.

Working with recombinant full-length proteins can be laborious and time consuming due to the fact that it is more complicated to produce these kinds of proteins [34]. To be able to ensure a native folding, they should be expressed in a system which makes it feasible e.g. mammalian cells [35]**.** Native proteins can also be used but it can be laborious to extract them from human tissue or body fluids.

If peptides are used as antigens, a length of about 10-20 amino acids are usually used to have functional linear epitopes. It is assumed that a peptide shorter than approximately 7 amino acids is insufficient to function as an epitope and if they are longer they could fold into a structure that is not found in the native protein [36]. Peptides have also successfully been used for epitope mapping of proteins. It was shown that a 10 amino acids long peptide was not enough to map the epitope but when 15 amino acids long peptides were used the epitope mapping was successful [37].

Recombinant proteins that are not full-length proteins are commonly used as antigens because they can be produced in large amounts, in e.g. *E.coli.,*

which can be used in the downstream process of the affinity reagent production. There are examples of both monoclonal, polyclonal and antibody fragments produced with recombinant proteins as antigen [38-42].

Antibodies

Antibodies are important components of the adaptive immune system. Antibodies circulates the body and have a function to eliminate microbes in the extra cellular fluids, but also to activate other components of the adaptive immune system [43].

Antibodies are glycoproteins secreted by specialized B lymphocytes, called plasma cells. They are also referred to as immunoglobulins (Ig) because they contain a common structural domain. Antibodies are built-up by four polypeptides, two identical copies of a heavy chain and two identical copies of a light chain, which are held together with disulfide and non-covalent bonds, see Figure 1. This structure is usually schematically presented like a Y-shaped molecule. Based on functional differences, antibodies can be further divided into two parts, the Fc–region and the Fab-region. The Fab domain is the arms of the Y-shape and contains the antigen binding sites of the antibody, that region is also referred to as the Fv region. The base of the Y is called the Fc region and the constant region of the Fc region is involved in binding between effector molecules in the complement system and more important to Fc receptors that are present on phagocytes and other immune system molecules [44]. Each variable region of the heavy chain and the light chain contains hyper variable regions, called complementarity determining regions (CDRs). These regions make it possible for the antibodies to adapt to almost any antigen that can appear in the body. In mammals there are five subclasses of Ig, classified depending on the conformation of their heavy chains (IgG, IgA, IgE, IgM and IgD). All have their own function in the immune system and can be used for different purposes in research. In the affinity research of today there are different types of antibodies available e.g. polyclonal, monoclonal and mono-specific antibodies but there are also fragments of antibodies used. Each of them have their strengths and weaknesses, but are useful in various applications.

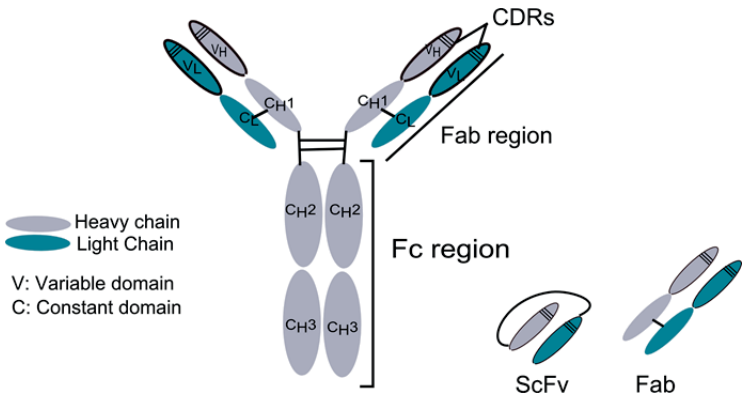

Figure 1. Schematic presentation of an antibody and the different regions it contains. Two different (ScFv and Fab) antibody fragments that are commonly recombinantly expressed are also shown.

Polyclonal Antibodies

Polyclonal antibodies are generally produced by immunizing an antigen to a host animal, usually mouse, rabbit or goat. The antigen is recognized as a foreign substance by the immune system of the host animal and the immune system starts to produce antibodies by specific B-cells. An antigen normally has several different epitopes that are recognized by the immune system and each specific B-cell will produce an antibody directed to one of the epitopes. This means that polyclonal antibodies will have a whole panel of different antibodies that recognize different parts of the antigen with different affinity. There is both an advantage and a disadvantage with a polyclonal immune response when antibodies are used as capturing tools. It is an advantage that different epitopes can be recognized by polyclonal antibodies. This can be used in protein microarrays due to the fact that the apparent affinity for polyclonal antibodies is higher because there are more binding sites and this will enhance the limit of detection [45]. The pool of polyclonal antibodies might be less sensitive to changes in the environment due to the fact that another epitope can be targeted by other antibody fractions of the pool if an epitope is destroyed [46]. The potential that lies in the polyclonal antibodies to recognize different types of epitope, both linear and structural epitopes [47] makes them useful in many kinds of applications [29, 30, 32, 34, 48-50]. Another advantage that makes the polyclonal antibodies useful is that they are relatively easy to produce compared to other types of binders. Already after a couple of months after the first immunization, it is possible to collect the antibodies from the animal [46]. Recognition of different epitopes could in some cases be a disadvantage because it is not exactly known which part of e.g. a protein that is targeted. It can be hard to distinguish an unspecific binding from correct ones. This fact makes it impossible to use polyclonal antibodies for *in vivo* diagnostics, but they are to a large extent used in *in vitro* diagnostics. There is also a reproducibility issue because there will be

different pools of antibodies produced in different immunizations even in the same animal. This will have the consequence, that more batch to batch validation has to be done for polyclonal antibodies than on a renewable resource of affinity reagent.

There are strategies that can be used to pull out specific antibodies from a polyclonal serum. Typically these antibodies are found in concentrations from 50 to 200 μg/mL [46]. For this purpose an affinity purification of the polyclonal serum on a column coupled with the antigen used for the immunization is done. Compared to monoclonal antibodies there will still be problems with reproducibility, since there is a need to produce new antibodies and their specificity may differ.

Monoclonal Antibodies

Monoclonal antibodies are often first choice when it comes to clinical diagnostics and research today due to the fact that they bind with one specificity to their targets. The principle, behind monoclonal antibodies, is that an antigen is immunized into a host animal, usually a mouse. So far, it is the same as the production of polyclonal antibodies, but in 1975 Köhler and Milstein published a paper where they describe how they fused activated, antibody producing cells with myeloma cells [31]. The hybridoma that is the outcome of this fusion is an immortal cell line that produces identical monoclonal antibodies. One advantages of this technology is that there will be identical antibodies that can be produced rapidly and reproducibly recognizing the same epitope. This can, however, be a disadvantage if the epitope that the antibodies are directed against is altered in a changing environment. Compared to polyclonal antibodies, the production of monoclonal antibodies are labor-intense and it can take more than a year to produce a functional monoclonal antibody. On the other hand there is a renewable resource when the cell line is established. More about the differences between monoclonal and polyclonal antibodies can be found in Lipman et al. [46]. There has been efforts to enhance the throughput of monoclonal antibodies by using robotics and antigen microarrays for screening [39].

Antibody Fragments and Scaffolds

Instead of using the whole antibodies, fragments, like single-chain variable (scFv) and fragment antigen binding (Fab), can be used. The Fab fragment was found as a product after enzymatic cleavage of the whole antibody with papain and pepsin and got the name simply because it was the antigen binging fragment [51]**.** Single-chain variables simply consist of the variable regions of the heavy and light chain and are coupled tighter with a linker [52]. In 1990 McCafferty et al. published an article where they used phage display to display and select antibody fragments [53]. Since then, there has been a lot of effort on developing phage display and there are a growing number of large libraries that consist of random combinations of antibody fragment

genes [40, 42, 54, 55]**.** In parallel with the further development of the phage display, other systems like ribosome display, cell surfaces display on Staphylococci and yeast displays have been developed and implemented in the generation of different kinds of affinity reagents [56, 57].

Antibody fragments are also produced in *E.coli* and already in 1988 there were two articles published that descried how functional antibody fragments could be produced in *E.coli* [58, 59]. This technology has since then been developed and is today widespread [60].

There are an increasing number of affinity reagents without immunoglobulin as template. One of them is a scaffold called, Affibody molecules. The origin of Affibody molecules are the B domain of protein A which is a *Stapylococcus aureus* protein and consists of 58 amino acids [61]. Another binder that has gained great interest is the Nanobody. The Nanobody derives from the unique heavy chain antibodies of camelids. Nanobodies are ~13-14 kD in size and are used in a wide range of biochemical applications [62, 63]. More about recombinant binders and their usability can be found in a recent review by Helma et al. [64].

Human Protein Atlas Project

The Human Protein Atlas project, founded 2003, is a large scale effort to produce antibodies against all human proteins [65-67]. The aim of the project is to produce at least one mono-specific antibody directed against each and every one of the 20.300 proteins. The HPA project would not have been possible without the sequencing of the human genome [12, 13]. All the work done in the project is based on the human genome sequence and the predictions that are made on which part of the genome actually will be expressed as a protein. Without that information, it would not be possible to create the protein fragments that are templates for the antibodies, which are used to generate protein profiles in human tissues and cells.

To be able to reach the goal, to produce antibodies against all proteins in the human body, a large scale production has been set up to produce recombinant protein fragments that are named Protein Epitope Signature Tag (PrEST) in the HPA project. To select a region of the full length protein, unique for the protein, an in house developed software is used. This web based program for antigen selection and visualization is based on data from Ensembl. In addition, there is also built in functions for prediction of transmembrane regions and predictions of solubility [3, 24]. A unique fragment of the protein consisting of 25-150 amino acids is selected. To avoid cross-reactivity the software selects protein sequences with as low homology as possible to other protein sequences in the human body. With this as a template, primers are designed in the software and the PrEST sequence is produced with RT-PCR with RNA as template. The sequence is ligated into a

vector that contains a hexa-histidine albumin binding protein (His_6ABP) gene sequence. The resulting His_6 tag on the protein is used for the purification of the PrESTs before immunization [68] and the ABP is used to enhance the immune response [69]. The vector construct is transformed into *E. coli.* The sequence of the clones is verified by Sanger sequencing [70]. After that, the *E.coli* colonies are cultured, the protein expression is induced and the resulting PrESTs are collected and purified [71, 72]. The PrESTs are immunized in rabbits and the polyclonal serum is collected. The first step in the purification procedure of the serum is a depletion column used to remove antibodies directed against the His_6-ABP tag. The serum is affinity purified on a second column with the corresponding PrEST coupled [41]. After Purification, there is what HPA refers to as a mono-specific antibody mixture directed against different epitopes on the PrEST [70]. The specificity of the antibodies is then verified by microarrays where the PrEST is used as target agent [41]. The antibodies are also tested in Western blot. In Immunohistochemistry (IHC) the localization of the corresponding protein is explored [41]. To get an even better view of the localization immunofluorescence (IF) is used for the subcellular localization [73]. In the beginning of 2015 the HPA project published an article with data covering more than 90% of the human putative protein-coding genes [17].

Validation of Affinity Reagents – an Important Factor in Proteomics

Why is it important to have validated affinity reagents to work with? First of all, it is a lot of money involved in affinity reagents. Since they are valuable and crucial for the assays, the purchaser needs to be sure that the reagent works. If not, there will be serious consequences as noted in a recent article by Monya Baker [74]. She refers to several publications in this matter and I will focus on some of them. In a study published by Begley and Ellis it was stated that out of 53 research papers based on landmark cancer research, 47 of them could not be reproduced [75]. This is a very sad number and although more efforts could possibly have been made to reproduce more of them, it is eye-catching, how large a proportion that failed. Even if this study is controversial it gives a hint of the tremendous challenge it is to establish reliable research tools within the proteomics field that can be trusted by the research community. This is one of the most crucial problems to solve today. In another study, 49 commercially available antibodies targeting 19 signal receptors were tested regarding their specificity. The conclusion was that the antibodies demonstrated a lack of selectivity that appeared to be the rule rather than the exception [76]. Egelhofer et al. showed that more than 25% of 246 histone-modification antibodies failed in specificity and even more

stunning, four of the antibodies were perfectly specific but to the wrong target [77]. On top of this, there is a number of other publications reporting on the bad performance of commercially available antibodies [25, 27, 66, 78]. It seems that there is quite a good awareness that there are a lot of antibodies that do not work and should not be on the market. There are however no consensus on how to get out of this mess, but there is an ongoing debate [26, 79-81]! Everyone agrees that more validation is needed to change what have been mentioned as the reproducibility crisis. This is where part of this thesis comes in, because **Paper I-V** is all about antibody validation. It sounds easy but antibody validation is not easy. First of all, it has to be stated that an antibody will/can work differently depending on the assay that it is used for. This is one part of the reproducibility crisis. In which context is the antibody used and what are the specifications from the manufacturer for that particular antibody?

If the affinity reagent bought is unspecific, much time is spent on investigating something that in the end shows to be of low value. Therefore, it is in the researcher's own interest that better and more similar validation strategies are set up for the affinity reagents that are commercially accessible. If there is no or limited information of the validation of the affinity reagent, the researcher has to put valuable time to investigate that the product that is bought works as it should. To speed up the growing market of affinity based diagnostics, there is a need for a more generally approved validation strategy for affinity reagents. It is important that this validation is not too time consuming and that the cost is not too high, which will limit the reagent to reach the market. Today there are many different ways to test the antibodies such as Western blot, immunoprecipitation and many more. Unfortunately there is no general agreement on what should be done and what kind of cut off that should be used to be able to state that an affinity reagent fulfills the criteria of being a validated and fully functional reagent. A first step could be if all producers of affinity reagents published all available data of the reagents.

What is then a specific antibody? It is not that easy to answer this question. James and Tawfik ask like this "Are specificity and promiscuity contradictory; or can a protein exhibit a promiscuous activity that is itself highly specific?" [82]. Their answer is that promiscuous activity does not necessarily need to be nonspecific. It is a matter of each antibody and the properties it has. Cross-reactivity involves binding of ligands that are similar or overlap the antibody's original antigen. Multi specificity on the other hand refers to the binding of distinctly different ligands, which can be the use of different set, combining site residues, or conformation differences of the same antibody. Promiscuous behavior has also been reported by Olsson et al. which they classify it as both specific and degenerate binding [83]. The promiscuous nature of antibodies further complicates the definition of specificity.

This picture of antibodies with a changing behavior is supported by findings in the HPA project. Boström et al. showed that HPA produced antibod-

ies can be used for IP but only about 50% of the tested polyclonal antibodies could capture the correct target even if all of them were tested on the antigen array [84, 85]. This clearly shows the need for verification of antibodies in the specific assay they are going to be used in. The largest difference between HPA antigens and native protein is that the former is denatured and shows linear epitopes. In native proteins those epitopes could be shielded or there is a structural epitope that this kind of antibodies does not recognize. IP can be used in non-denaturing [86] or denaturing [87] mode. Running them both when evaluating antibodies in combination with MS would probably improve the outcome. In MS analysis the folded or unfolded protein does not matter which means that if it is only a matter of quantification with MS a denaturized protein IP could be used. Schwenk et al. have shown that heat treatment of plasma sample improved detection of components of the complement system using HPA antibodies [88]. Their explanation is that epitopes will be exposed at elevated temperature thereby linear epitopes will be accessible for the HPA antibody. In a large study, Marcon at al. set up a mass spectrometry–based standardized protocol characterizing antibody selectivity and specificity for use in IP that was evaluated in five independent laboratories [89]. This kind of large scale initiative is a good example of what has to be done to improve the outcome of affinity based research and is also a nice example on how affinity and mass spectrometry-based methods can be used in combination to improve research tools within proteomics.

Protein Microarrays - Applications

A microarray is an ordered structure of samples such as DNA, proteins or other samples of interest. This technique took off during the 1990s with the evolvement of the DNA microarrays [90, 91]. The principles of miniaturized and parallelized immunoassay were already described in the late 1980s by Roger Ekins [92]. However, it was not until the late 1990s that protein microarrays were developed [93-96]. At that time DNA microarrays and the tools needed to produce microarrays in a high-throughput fashion were established. The principle is to miniaturize and parallelize the assay to get as much information as possible in a single run. However, it was challenging to adapt into protein microarrays, since proteins are far more complex to work with than DNA. One of the largest limitations was and still is the production of proteins used for the production of the protein arrays. Today, protein microarrays are established and used in many different applications [56, 97]. There are different names of them. Here I classify them as antigen arrays, antibody arrays and reverse-phase microarrays, See Figure 2.

Antigen Microarray

An antigen microarray consists of immobilized antigens on a solid support. Antigen microarrays were used as a validation tool of polyclonal and monoclonal antibodies as well as scFv in **Paper I-IV**. It was shown that antigen microarrays can be a valuable tool for validation when new affinity reagents are produced. The specificity of an affinity binder is important, in order to be confident that the molecule that is detected is the correct target. It has recently also been nicely presented that immunoprecipitation (IP) assay coupled to MS could be a useful tool for verification of the specificity of the produced antibodies [84, 89].

Antigen microarrays have also been used for diagnosis of allergy and then different allergens are printed on the array [98, 99].

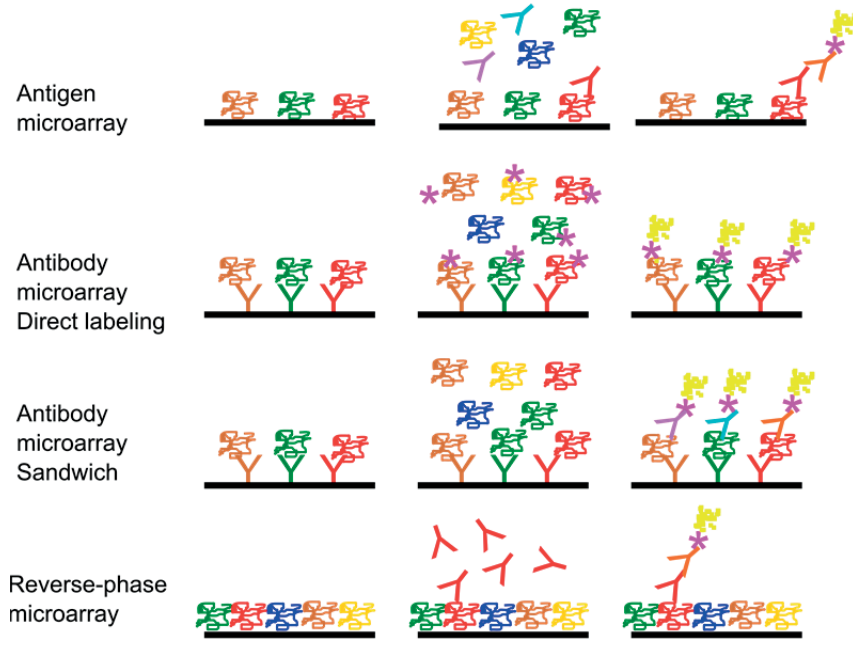

Figure 2. Three types of protein microarrays with printed array to the left, incubation in the middle and detection to the right. The presented set up is the one used in this thesis and related publications.

Another application for antigen microarrays are screening for human autoantibodies. There is a growing need to get a deeper understanding of different autoimmune diseases and here antigen microarrays could be an important tool to give new information and new targets that could be used in both diagnosis and therapy of these diseases [100-103]. There have been studies with antigen microarrays as a tool for identification of the immunodominant

pathogen in different infectious diseases [104, 105]. More about different antigen microarray applications can be found in Ayoglu et al. [56].

Antibody Microarray

Antibody arrays can be produced with polyclonal, monoclonal, single chain variable (scFv), Fab, Affibody molecules or other kinds of affinity binders. Antibody microarrays are together with antigen microarrays the most widely used protein microarrays. There is an example where small size arrays are used in a diagnostic application. One example is diagnosis of acute coronary syndrome (ACS) [106]. More about the status of antibody based microarrays for clinical applications can be found in theses review articles [107, 108]. A limiting factor is the expense and lack of high quality antibodies. This is indicated of the size of the produced and published studies with antibody microarrays. It is usually from ten and up to a couple of hundred antibodies that are included. One of the larger antibody arrays, produced by Schröder et al. contained 810 antibodies directed against 741 different proteins [109]. There are already initiatives taken to use the arrays for large scale protein profiling, even if the one referred to here is based on a bead based assay [88]**.** This kind of initiative will open up a new field that can be used in parallel with e.g. mass spectrometry protein profiling. This will hopefully give many new biomarker candidates and potential therapeutic targets. Antibody microarrays and MS can also be used for orthogonal verification of studies. There is a lot of potential in using antibody arrays both as a diagnostic tool and a biotechnological tool. The antibodies and the produced microarrays have to be thoroughly validated before large scale use.

Peptide microarrays have in the last years been introduced as a complement to antigen microarrays. Peptide arrays have been used epitope mapping of proteins with promising results [110, 111].

Reverse-phase Microarrays

A reverse-phase microarray is constructed by spotting a complex sample e.g. human tissue lysate or serum/plasma on a solid support. The strength of doing this is that thousands of patient samples can be screened in a single run at equal conditions [112]. Due to the fact that the amount of sample deposit is small, there is not much patient sample needed for the analysis. This is, however, one of the drawbacks with these kinds of microarrays because the number of molecules of each protein in the sample will be very low, it might be below 100 molecules [56]. In a study done by Janzi et al. with a serum microarray targeting IgA, the sensitivity was in the lower µg/ml range [112]**.** This shows that the sensitivity today is not good enough for the low abundant proteins (concentrations below µg/ml) in these kinds of studies. Although it can still be of interest due to the fact that thousands of samples can

be tested at the same time. For medium to high abundant proteins this could be useful. Still there are possibilities to enhance the signal and there have been some studies to do this [113]. In the future there might be new amplification techniques that could be used to increase the sensitivity for reverse-phase microarrays and then they could be a very good tool for large scale screenings of large patient cohorts.

Production of Protein Microarrays

The protein microarray production is based on the printing proteins with a robot and immobilized the proteins on a solid surface. Microscope slides are usually used, with an adsorbent or a chemically activated surface. This implies that there is a small amount of protein solutions that is deposited either by a drop or by physical contact with a steel pin. The amount of material that is deposited is usually 100-400 pL. This will result in a spot size of about 100-300 micrometer on the solid support. Tens of thousands of spots can be placed within one slide. The possibility to print many different proteins in one slide is one of the main advantages with planar protein microarrays. Screening a lot of proteins with many different patient samples is feasible, but the throughput will be rather low.

Table 1. Some surfaces that can be used as solid support in protein microarrays.

Surface chemistry	**Binding mode/ reactive groups**
Epoxy	Covalent/ amino, thiol or hydroxyl groups
Aldehyde	Covalent/ amino groups
Poly-L-Lysine	Adsorption, ionic interaction
Aminosilane	Adsorption, ionic interaction
Amine	Covalent/ amino groups
NHS	Covalent/ amino groups
PVDF	Adsorption
Nitrocellulose	Adsorption
Polystyrene	Adsorption/polar groups
Polyacrylamide gel	Adsorption, interactions with the 3-D matrix
Biotin immobilized in polycarboxylate hydrogel	Streptavidin/avidin/ neutravidin
Polycarboxylate hydrogel, NHS activated	Adsorption, amino groups and interactions with the 3-D matrix
Ni^{2+} ions complexed in polycarboxylate Hydrogel	Poly histidine tag
Protein A polycarboxylate hydrogel	Human immunoglobulin proteins
Streptavidin polycarboxylate hydrogel	Biotin

A very important component when producing protein microarrays is the printing robot, which is a device, precisely delivering small amount of the sample onto a solid support. There are two main methods of printing microarrays and those are, contact and non-contact printing. A review of different printing technologies has recently been published by Romanov et al. [114]

There is a large diversity of solid supports that can be used for protein microarrays on the market [96, 115, 116]. In Table 1, different surfaces are listed that can be used and the binding mode that is used for the immobilization.

Blocking the surfaces that are not immobilized with protein is another important issue when working with protein microarrays. This can be tricky and a lot of testing might be needed to find a good blocking solution. Usually this is done with some kind of protein blocking. The most common used is bovine serum albumin (BSA) or milk or a combination of them. There is also a growing commercial market of blocking solutions.

Incubation, Scanning and Data Evaluation

There are many different detection systems that can be used, but in protein microarrays fluorescence based are the most frequently used [117]. Some research groups are using chemiluminescent detection [118].

There are two different ways of labeling samples that are being tested: direct labeling or indirect labeling. See Figure 2. Indirect labeling, called sandwich assay, is an approach where the proteins in a complex sample like plasma are captured on an antibody array, then a second complementary fluorophore conjugated antibody is introduced for detection [119]. Direct labeling means that the sample is labeled with e.g. a fluorophore and then incubated on the slide [120]. In antigen microarrays the detection is usually done with a fluorophore conjugated secondary antibody which is specific for the host animal in which the primary antibody is produced.

There have been studies to compare different labeling strategies to find out advantages and disadvantages with the different approaches [121, 122]

The sample of interest (labeled or not), are applied to the slide and incubated for 1h up to 24 h. After that, the unbound material is washed away. If the sample is not directly labeled with a fluorophore, a secondary antibody e.g. an anti-rabbit/anti-human antibody label with a fluorophore or other kind of detectable reagent is put on and incubated, usually 1 h. When this is done the slide is washed again to wash away unbound secondary antibody and after that the slide is dried and scanned.

The evaluation of the scanned slide is done with an image analyzing software, which translates the fluorescent signals into numbers that can be analyzed using different kinds of statistical methods, often in a statistical program like R or equal.

Bead Based Protein Microarrays

In a bead based system the solid support is color-coded microspheres that can be identified with a flow cytometry based system [56, 95]. In **Paper III** a bead based set up was compared to a planar antigen microarray. There are 500 unique color-coded microspheres available, which is the upper limitation of how many analytes can be tested. Each of these beads can be coupled to an antibody or antigen in suspension and after coupling the different beads can be mixed together. This is very flexible due to the fact that when the beads are coupled there are no limitations how they can be mixed together. In contrast to planar arrays the whole procedure is done in suspension which means that there is no drying that can make the coupled molecule inactivated or denatured. The suspension setup gives better possibilities to make the whole process automated by robots. Compared to planar arrays the bead arrays are faster when analyzing many different patient samples because it is only adding more wells in the microplate and there is no image analysis. If slides are used, many slides have to be included and there is also a scanning and image analysis.

The procedure starts with coupling the beads, a specified amount of antibodies or antigens are added to a specified amount of beads. Then beads are mixed together to create a microarray in suspension. Then sample is incubated with the beads. In bead based system the detection is done by fluorescent dyes. More details about bead coupling and sample labeling are found in two articles by Schwenk et al. [88, 123]. When the incubation is done the incubated beads are detected in a flow cytometry based system equipped with two lasers. One detects the identity of the bead and the other the intensity of the signal from sample. In the output file, the mean intensity of all the beads with the same identity is reported as well as how many of that specific bead identity that are counted. More than one bead is always counted, at least 50 beads/well of the different bead identities is usually included. The output data is a mean value based of at least 50 different beads and the signal generated from them. Looking into a user perspective, this system is much more user-friendly and it is much more flexible because the different beads can be mixed in any combination while the planar array have a fixed pattern even if many spots can be included but then the cost can be an issue.

Gyrolab Technology

The Gyrolab technology is based on a parallel nanoliter microfluidic approach and was used in **Paper II**. The platform builds on a special type of compact disc (CD) with integrated micro capillaries minimizing the sample consumption. The transport of the fluid is controlled by capillary action, centrifugal force and hydrophobic barriers [124]. An on-column detection is

performed, using a scanning confocal laser-induced fluorescent detector, integrated in the workstation. The CD contains 112 reaction chambers that can be used for separate analysis. The etched chambers and capillaries are covered with a lid to avoid evaporation of the small volume used. The sample is loaded onto the CD via an inlet on top of the CD. A needle integrated in the workstation is used for dispensing the sample from a 96-well plate. The sample is then loaded through capillary forces. To be able to have a well-defined volume, a hydrophobic barrier is used to fill up 20 or 200 nL chambers. Centrifugal force is used to get the sample pass these barriers.

The data that is collected from the Gyrolab workstation is processed with the software Gyrolab viewer. The integrated values of the detected signals can be presented as two-dimensional curves based on the sum of fluorescence intensity or as a 3D presentation of the peak based on the total integrated volume [125].

3. Proteomics Through Mass Spectrometry

Mass spectrometry is today the most widespread and versatile technology used to conduct proteomics research. The technique is based on detection of molecular ions in gas phase and there is a wide variety of instruments. The basal set up of an instrument consists of three parts: the ion source and optics, the mass analyzer and the detector. The ion source enables the formation of ions into gas phase either from a liquid sample or from a solid surface. In the analyzer, ions are selected and separated based on their mass-to-charge ratio (m/z) and finally ions are detected. There is a wide range of ionization sources, analyzers and detectors and more information of them can be found in e.g. de Hoffmann and Stroobant [126]. There are two ionization techniques, invented in the late 1980s, that have made the fast development in mass spectrometry based proteomics possible. These two are Matrix Assisted Laser Desorption Ionization (MALDI) [127, 128] and Electrospray ionization (ESI) [129, 130].

Table 2. Some properties of mass analyzers used in proteomic studies [126].

Mass analyser	Mass limit (Th)	Resolution	Accuracy (ppm)
Quadrupole	4000	2000	100
Ion trap	6000	4000	100
TOF	>1000 000	5 000	200
TOF reflectron	10 000	20 000	10
FTICR	30 000	500 000	<5
Orbitrap	50 000	280 000	<5

In proteomics research usually one of the following mass analyzers are used: quadrupole (Q) [131], ion trap [132], time-of-flight (TOF) [133], Fourier transform ion cyclotron resonance (FTICR) [134] or orbitrap [135]. Mass spectrometers have different mass limits (unit Thomsson (Th)) i.e. the maximum mass-to-charge ration that can be analyzed. There are also differences in resolution, which is the ability to distinguish between two peaks of slightly different mass-to-charge ratios. Finally, the accuracy of various instruments differs. All these parameters are important for the ability to detect proteins. The purpose of study and availability of instrument determines which instrument is the best to use. High resolution can be very good but that will generally slow down the data acquisition and thereby proteins can be missed. Table 2 lists the most frequently used analyzers in proteomics

together with some of their properties. The most common in proteomics is, however, to use different combinations of analyzers, named hybrid instruments [126, 136]. In this thesis, ESI in combination with either a LTQ Orbitrap Velos Pro ETD [137] (**Paper V-VII**) or a Q Exactive Plus Orbitrap [138] (**Paper VI**) have been used. Therefore, these will be more thoroughly described.

Liquid Chromatography and Electrospray Ionization

A modern mass spectrometer is fairly good to separate different ions that are introduced. However, in proteomics the sample is usually very complex with a large number of proteins that should be detected. Therefore, a chromatographic step is introduced prior to MS analysis. The miniaturization of liquid chromatographic systems has been beneficial for MS [139]. The sample consumption has decreased and it has been possible for on-line combinations of LC and ESI source. The most commonly used chromatographic material in MS-based proteomics is reverse phase [140]. Other materials used are ion exchange (IEX), hydrophilic-interaction chromatography (HILIC). If those combinations do not solve the separation issue, capillary electrophoresis can be used [141]. Reverse-phase columns are usually coupled to a HPLC or Ultra Performance Liquid Chromatography (UPLC) systems. A great advantage using reverse phase is that the buffers used are compatible with ESI [140]. There are a lot that can be changed to optimize the performance of the column. Things that can be modified are e.g. dimensions of the column, the particle size or the column can be heated [140].

For peptides, an octadecyl carbon chain (C18)-bonded silica (reverse phase) can be used and C18 is used in this thesis. There are other variants of the stationary phase e.g. C8, C4 and diphenyl. To be sure not to contaminate the analytical column it is common to use a pre-column as a protective cleaning step.

In 1988 John Fenn and coworkers for the first time showed that they could transfer large molecules into gas-phase without breaking them [142]. Since then this is an invaluable tool in bio molecular research. ESI together with MALDI are soft ionization techniques which means that there is very little fragmentation of proteins/peptides during the ionization process. It is easy to connect ESI with liquid chromatography (LC). ESI was also the ion source used in the mass spectrometry analysis preformed for this thesis.

Figure 3 gives a schematic picture of the electrospray ionization. To generate an electrospray, a current potential is applied, typically 1-3 kV, between the spray emitter and the inlet of the mass spectrometer. This induces a charge accumulation at the liquid surface located at the end of the capillary and the so-called Taylor cone is formed. This will break to form highly charged droplets. These droplets start to evaporate and are electrically accel-

erated towards the heated inlet of the mass spectrometer. During that process, the analytes (i.e. peptides) are released as gas-phase ions [126, 130]. How this works is not fully understood, but two models are proposed for the ionization process: the ion evaporation model (IEM) [130, 143] and the charge residue model (CRM) [130, 144]. The IEM suggests that when droplets reach a certain radius (~10 nm), the field strength at the surface of the droplet becomes large enough and direct ion emission becomes possible. The CRM model suggests that electrospray droplets undergo evaporation and fission cycles which will lead to droplets that contain on average one analyte or less. It is stressed that these are models and that not everything can be explained by a model. For large molecules (e.g. peptides/proteins) much of the CRM model is applicable [130].

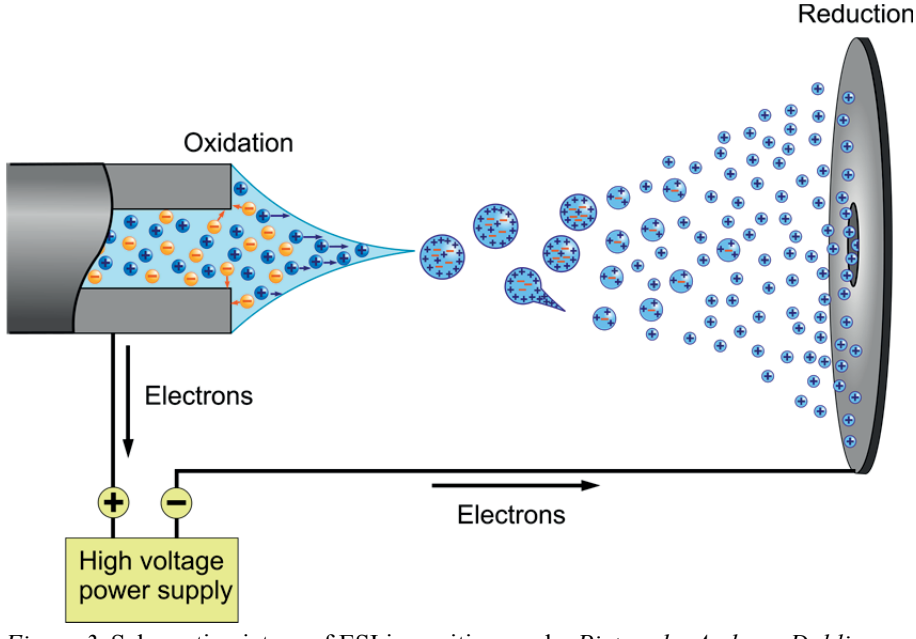

Figure 3. Schematic picture of ESI in positive mode. *Picture by Andreas Dahlin*.

Orbitrap Mass Spectrometers

The Orbitrap mass analyzer was first presented 2000 by the inventor Alexander Makarov. The construction consists of an outer barrel-like electrode and a coaxial inner spindle-like electrode which traps ions in an orbital motion around the spindle. Fourier transformation of the frequency signal is then used to convert the image current from the trapped ions to a mass spectrum [135]. The principle behind the Orbitrap will result in a mass spectrometer with very high resolution almost comparable to an FTICR. Resolution of 280 000 at 200 m/z is achieved on the Q Exactive plus instrument. It should although be noted that the resolution decreases with the square root of the

m/z in Orbitrap analyzers [138]. One of the big advantages compared to an FTICR is that an Orbitrap does not need extensive cooling as an FTICR. The Orbitrap has a large dynamic range and deliverers high mass accuracy [137].

In a mass spectrometer the detection can be of the introduced peptides and that is referred to as MS1. The mass spectrometer can in many cases also fragment the peptides and then detect the fragments and that is referred to as tandem MS (MS/MS or MS2). By introducing a fragmentation step, a more reliable assignment of the introduced peptides will be achieved, since information on the peptide sequence is also obtained. For example, two peptides can have very similar or the same masses, but different amino acid order. These peptides cannot be distinguished if the identification is based on MS1. It is possible to do fragmentation in the Orbitrap which is called all-ion fragmentation [145]. The problem is that no separation of the ions can be done in the Orbitrap. This is solved by hybrid instruments and in this thesis, two hybrid instruments with an Orbitrap analyzer have been used (See Figure 4).

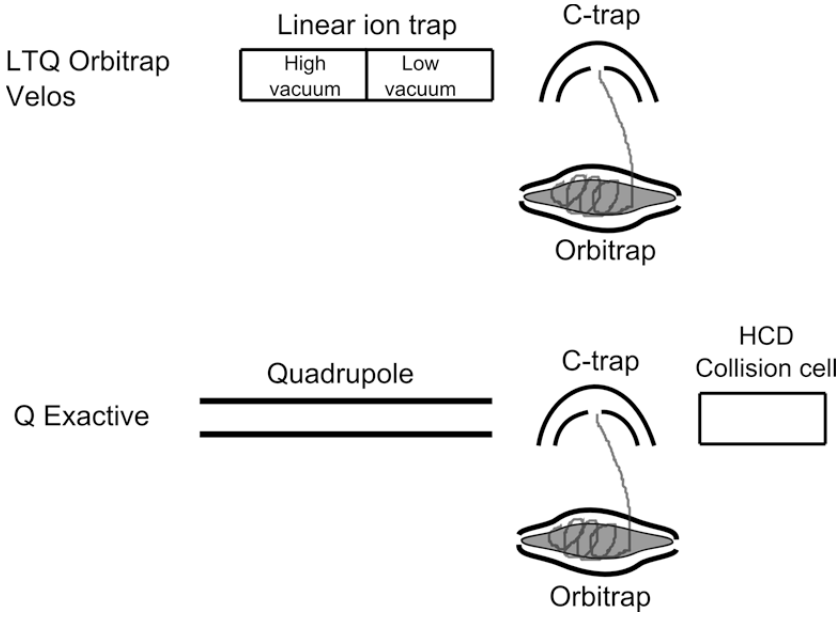

Figure 4. Schematic view of two hybrid instruments with Orbitrap mass analyzers.

The LTQ Orbitrap Velos Pro ETD has a linear ion trap which offers very high sensitivity, fast analysis and has the capability to do fragmentations in multiple levels, usually referred to as MS^n. The linear ion trap consists of a high vacuum part for fragmentation and a low vacuum part for analysis. The fragmentation in the linear ion trap is performed with collision induced dissociation (CID). In the Velos Pro ETD version, it is also possible to achieve fragmentation by Higher-energy Collisional Dissociation (HCD) or Electron

Transfer Dissociation (ETD). The next important part in this hybrid instrument is the C-trap which is a bent RF-only quadrupole that is capable to accumulate and store ions. The last part is the Orbitrap that is filled via pulses of ions accumulated in the C-trap. The Orbitrap and the ion trap can be operated separately or in combination. The Orbitrap has much higher resolution but has due to that slower acquisition. The aim of the study determines the best way to operate the instrument. Speed and resolution are important factors but are difficult to combine in a ultimate way and compromises have to be done [137].

In a Q Exactive Plus, the linear ion trap is exchanged with a quadrupole which gives a very fast mass selection and in combination with HCD the detection and fragmentation occur more or less simultaneously. There is no possibility to do MS^n fragmentation in a Q Exactive Plus. The quadrupole cannot work by itself which means that this instrument always will work as a unit. Another feature on the Q Exactive instrument is that multiplexing up to 10 times can be done both in MS and MS/MS. In multiplexing, precursors or fragments are selected from an inclusions list, stored and finally analyzed simultaneously in the Orbitrap. This is possible due to the storage capacity in the HCD cell and the fast selection in the quadrupole [138].

Q Exactive is an excellent instrument for targeted analysis and was the instrument chosen for the analysis in **Paper VI**.

In a Q Exactive Plus, the linear ion trap is exchanged with a quadrupole which gives a very fast mass selection and in combination with HCD the detection and fragmentation occur more or less simultaneously. There is no possibility to do MS^n fragmentation in a Q Exactive Plus. The quadrupole cannot work by itself which means that this instrument always will work as a unit. Another feature on the Q Exactive instrument is that multiplexing up to 10 times can be done both in MS and MS/MS. In multiplexing, precursors or fragments are selected from an inclusions list, stored and finally analyzed simultaneously in the Orbitrap. This is possible due to the storage capacity in the HCD cell and the fast selection in the quadrupole [138].

Q Exactive is an excellent instrument for targeted analysis and was the instrument chosen for the analysis in **Paper VI**.

Sample Preparation

One crucial step in successful mass spectrometry proteomics is sample preparation. Proteins can be small like insulin composed of only 51 amino acids up to Titin which is the largest known protein with up to ~33,000 amino acids. The sample preparation depends on the purpose of the study and how the sample will be introduced to the mass spectrometer. Here some of the most common ways of sample preparation for introduction with ESI will be described. A protein of interest can be found in many kinds of sample matri-

ces such as body fluids (plasma, sera, urine etc.). The sample can also be any kind of tissue, but then some kind of extraction has to be done to have the proteins in liquid form before being introduced into the mass spectrometer. There are basically three different ways of analyzing proteins and those are Bottom Up/Shotgun [8], Top Down [146], and Middle Down [147] proteomics (See Figure 5).

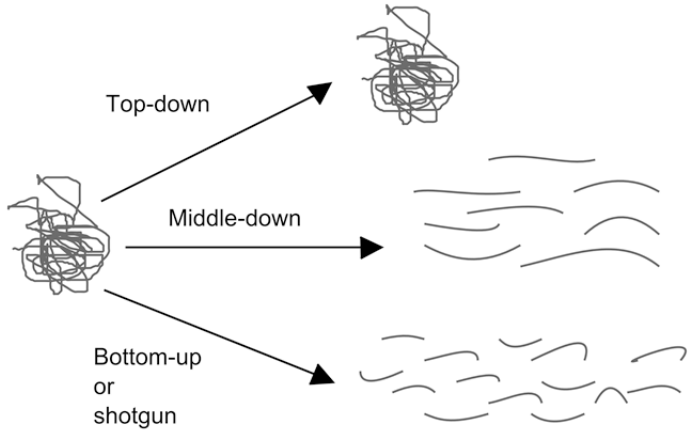

Figure 5. Three different ways of protein analysis using mass spectrometry. The protein can be either analyzed intact (top-down) or it can be proteolytic digested. In middle-down the protein is digested to give larger fragments. In bottom-up the sample is digested with e.g. trypsin that generates relatively short peptides that are analyzed.

Bottom Up or shotgun proteomics is the outstanding most used mass spectrometry based method for proteomic studies and is used in **Paper V and VI**. This method is based on chemical or enzymatic digestion of proteins before introduction into the mass spectrometer. The resulting peptides are detected and then an extensive data processing is needed to identify and connect the peptide with a protein. The most common way to digest proteins in front of a shotgun analysis is to use trypsin, which is used in **Paper V-VII**. It has been pointed out that the major focus on tryptic digestion hampers the proteomics research [148]. An example is that 257 human proteins (isoforms not included) do not produce any tryptic peptides [16]. An overwhelming part of all shotgun studies have been conducted with the aid of trypsin, but there are a lot of other proteases that also could be used for digestion. The most used proteases, except for trypsin, are Lys-C, chymotrypsin, Glu-C and pepsin [148]. Digestion can either be done in-solution, in-gel or on-filter.

A small protein like Insulin would be possible to analyze using a mass spectrometry method called Top Down where the intact proteins are injected to the mass spectrometer [146]. It is challenging to conduct Top Down pro-

teomics and it can be really difficult to analyze many proteins of different sizes in a complex mixture like plasma compared to analyze a small pure protein. Top Down can be a good method for PTM analysis, but so far this application is not that much used due to the complexity to run it. Top Down lag behind Bottom up when it comes to throughput, sensitivity and proteome coverage.

Middle Down is also based on digestion but, with a protease producing larger peptides. This method is used as a complement to Top Down analyzing larger proteins from antibodies (~150 kDa) and above. These kinds of proteins can be really difficult to analyze using Top Down and a lot of information about structure and post translational modifications could be hidden analyzing the sample with Shotgun [147].

Data Handling

After analysis there are a large amount of spectra that have been collected. To be able to identify the amino acid sequence of the mass spectrometer analyzed peptides, the amino acid sequences of the proteins are downloaded from a protein database, like Uniprot. If the sequence is not known, the identification step is much more complicated and this is why large scale sequencing of genomes has had a big impact on the development of proteomics. If the amino acid sequence is unknown, the much more time consuming *de novo* peptide sequencing must be conducted [149]. When the sequences are downloaded an *in silico* digestion is done and then a comparison with the mass spectrometry data can be performed. To do this there is a growing number of search algorithms that can be used, such as SEQUEST [150], MASCOT [151], MS Amanda [152], X! Tandem [153], Andromeda [154], MS-GF+ [155] and Byonic [156]. To increase the number of correctly identified peptides and thereby decrease the false discoveries, different algorithms like Percolator [157] or peptide decoy peptide spectrum match (PSM) validation [158] are used. The first one is a self-learning iterative process to validate PSMs as correct or incorrect. The peptide decoy does this by just reversing the peptide sequence and use this false "decoy" sequence as comparison to find incorrect results in e.g. a SEQUEST search. This falls back on what is crucial in both mass spectrometry based as well as affinity based proteomics assays, i.e. sensitivity and specificity. *Sensitivity* is defined as the ability to deliver a positive test result with high probability i.e. low number of false negative answers. High *specificity* in a method is defined as the ability to measure the true negative i.e. low number of false positive.

In the last decade there have also popped up many different databases for searching and sharing proteomics data. These may originate from different human proteomics initiatives such as the Human Proteome Project [159], the Human Protein Atlas project [17], ProteomicsDB [16], Human Proteome

Map [15]. Then there are also more general search tools for proteomics research such as neXtProt [160], NCBI [161] and the most used database for this thesis, UniProt [14]. There are also many different databases exclusively for mass spectrometry based proteomics data. In these databases you are encourage to deposit your data set and you are also free to reuse the data for new searches. Examples of these kind of databases are GPMDB [162], PaxDb [163], PeptideAtlas [164], PRIDE [165]. ProteomeXchange [166] is a consortium started with the intention to make it easier to share mass spectrometry data. They support infrastructure to make the upload of data easier and both shotgun as well as targeted proteomics data can be shared via this portal. This is a very good initiative due to the fact that the number of different databases will make it very time consuming to get a full picture of the proteomics research today. Even if several of the mentioned databases store data from different species and most of them have references to each other, there is today no single source that can be used for a complete proteomics data search. Within this area there is a lot to do in order to facilitate effective research. If one could use only one portal for protein search and comparison between species it would contribute to the whole research community. Another problem related to this is all the different file formats that exist and the lack of a common consensus of how the data analysis including the statistical analysis should be conducted. This lack makes it virtually impossible to make direct comparisons between studies.

Mass Spectrometry-based Quantitative Proteomics

Quantification can be conducted in two ways, either as relative comparison between samples or as an absolute determination of the protein amount in the sample. Quantification in proteomics is not new but there has been a tremendous development of mass spectrometry-based proteomics the last 15 years. In 2005 Ong and Mann published the review article “Mass spectrometry-based proteomics turns quantitative” and they could not have been more correct [167]. The last decade there has been an explosion of publications regarding mass spectrometry-based quantification. This is a very expansive and diverse area of proteomics, showing the importance of having quantitative methods. There are a very large spectrum of methods described and for the interested reader there are very good review articles covering most of this field [168-173]. In this thesis one of the studies (**Paper VI**) is a contribution to targeted absolute quantification. In Figure 6 workflows for the different approaches are presented. There are differences in when and if a labeling procedure is introduced. If the labeling is done in the cell or tissue, the sample preparation is the same for all samples, which will limit the variation between the samples. The later the samples are combined, the bigger

impact of sample preparation and variance between the samples. This will increase the risk of a less accurate quantification.

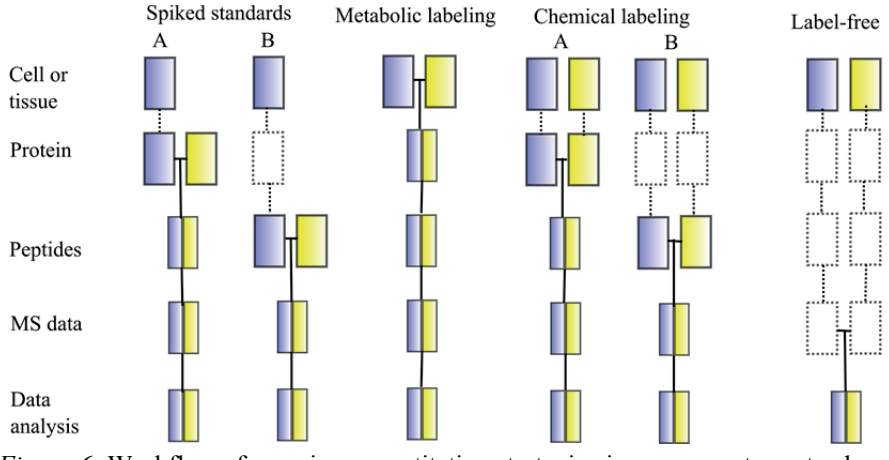

Figure 6. Workflows for various quantitative strategies in mass spectrometry-based proteomics. Blue and yellow represent samples to compare. Dashed lines shows where variation can be introduced. Horizontal line marks when samples are combined. (A) indicates that full-length proteins are used as standard and (B) indicates that peptides are used as standard. (Adapted from [167, 168])

Label-free Quantification

Label-free quantification is widely used due to the fact that it does not include any extra cost of reagents and it is easy to include many samples for comparison. Label-free quantification consists of two different strategies for relative quantification of proteins in different samples. The methods are either based on the number of peptide spectrum matches (PSMs, spectrum count) [174] or peptide signal intensity (or peak area) [175, 176]. Spectrum count has been questioned because the quantification is based on counting the number of acquired spectra rather than measurements of physical data [168]. Methods based on peptide signal intensity need specialized software for data handling which makes this approach more laborious. Label-free quantification is a cost effective method to get quantitative data from large data sets. Studies have shown that label-free methods compared to quantitative labeling methods have the larger dynamic range and higher proteome coverage [177-179]. In a discovery phase of a project, label-free quantification it can be useful to select a smaller number of candidates that could be further studied. There are also label-free methods for absolute quantification. To be able to do this, a different kind of data normalization is needed. Different methods and their advantages and disadvantages have been reviewed recently [168, 170].

Stable Isotope Labeling Strategies for Relative Quantification

Stable isotope labeling of amino acids is another approach to get quantitative data out of a mass spectrometry analysis. These techniques give added value compared to label-free quantification. The principle behind this is that there is a "light" unlabeled form and a "heavy" labeled form of the peptides in the sample that will be analyzed. The labeling is done with non-radioactive isotope labels such as ^{13}C, ^{15}N, ^{2}H and ^{18}O, with the two first most used. Deuterated peptides can separate in reverse-phase chromatography and therefore ^{13}C and ^{15}N are preferred [180]. The ionization efficiency of them is the same for both forms. This will give the same response signal, except a mass shift that can be detected in the mass spectrometer. There are several methods available that are based on stable isotope quantification in proteomics. There are four categories of which mainly the three first are used today. They can be classified as (I) metabolic incorporation of the labeling, (II) chemical introduction of an isotopic tag into proteins or peptides, (III) spiking in an isotopically labeled analog, (IV) introduction (^{18}O) through an enzyme during the digestion of the protein [167]. To get a good separation of the analogs in a mass spectrometer, a mass shift of at least 3- or 4-Da is needed in order to minimize quantitative errors due to isotopic overlap. In this thesis the spiking of isotopically labeled analog have been used and will be further discussed below.

In metabolic labeling, the isotopic labeling is introduced to whole cells through the growth medium. Ong and co-workers presented 2002, the method stable isotope labeling of amino acids in cell culture (SILAC) as a relative quantitative method [181]. The most common is to label arginine and lysine (^{13}C, ^{15}N), which will ensure that all tryptic peptides will have at least one labeled amino acid [168]. Chen et al. has recently published a thorough review of SILAC [182].

Another way of labeling is to use some kind of chemical labeling. The first method that was presented using isotopic labeling was isotope coded affinity tags (ICAT) [183]. This method is based on labeling cysteines. Dimethyl labeling is a cheap and fast labeling method [184] that can also be multiplexed to 3-plex [185]. The last two chemical labeling methods that will be mentioned are isobaric tags for relative and absolute quantification (iTRAQ) [186] and tandem mass tag (TMT) [187]. These two methods are also performed at peptide level and primary amine residues are the labeling targets. These methods have much better multiplexing abilities compared to other chemical labeling methods, iTRAQ can be multiplex up to 8-plex [188] and TMT up to 10-plex [189]. All of mentioned methods need sophisticated software for the data analysis. The analysis can be very complex and that's also a limitation for this kind of methods.

Stable Isotope Labeled Standards for Absolute Quantification

Absolute quantification of peptides (and thereby proteins) can be achieved by addition of stable isotope-labeled reference standards with known absolute quantity. [190, 191]. A comparison of the MS signal intensities of the endogenous peptides and the corresponding standard peptides gives the absolute concentration of the protein of interest.

There are several kinds of standards that can be used, both full-length proteins, protein fragments and peptides. In 2003, Absolute QUAntification (AQUA) [191] was presented and after that more methods have followed based on stable isotope labeling. AQUA´s are synthetic heavy isotope labeled peptides that are commercially available from several different vendors. Although quite expensive this reagents is widely used and particularly in combination with targeted approaches such as SRM, see targeted absolute quantification [168]. One problem with AQUA peptides is that they are spiked into the sample after the digestion and thereby the digestion efficiency could affect the quantification. Another approach is QconCAT [192], which uses recombinant DNA techniques for construction of concatenated peptides. These constructs can then be expressed in stable isotope-labeled form in a suitable host e.g. *E. coli*.

To be able to include standards as early as possible, there are also methods that use full-length proteins, such methods are PSAQ (for protein standard absolute quantification) [193], absolute SILAC [194], FlexiQuant (for Full-Length Expressed Stable Isotope-labeled Proteins for Quantification) [195] and RISQ (for recombinant isotope labeled and selenium quantified) [196]. In this thesis, a recently launched type of regent called QPrEST [197] was tested in a targeted approach (**Paper VI**). The very well characterized PrESTs from the HPA project were expressed with heavy labeled (^{13}C, ^{15}N) arginine and lysine. QPrESTs have been used in a SILAC approach to determine the protein copy numbers of proteins in HeLa cells [197] and murine platelets [198]. QPrESTs are not full-length proteins and are therefore not suitable to be added before protein fractionation as full-length proteins can be. On the other hand, they are easier to express than full-length proteins and less expensive. Another advantage of QPrESTs is the large library of well validated constructs, covering > 80 % of the human protein-coding genes. This means that the construct is already validated and ready to use. Which is not the case for QconCAT where the construct has to be constructed. Added to that, the QPrESTs have a much larger similarity to the full-length protein than a QconCAT construct, which will give a more comparable digestion. In **Paper VI**, we showed the possibility to use QPrESTs for other animals, not restricted to human samples. All presented standards have their advantages and disadvantages and which of them that will be used as “golden standard” is still unclear and at least some of them will be used in parallel due to different properties useful in different kinds of quantification studies.

Targeted Absolute Quantification

Mass spectrometry quantification of proteins is today based on two acquisition modes. In a shotgun set up, the instrument is operated in a data-dependent acquisition mode (DDA). This means that the precursors (peptides) with the highest intensities are chosen for fragmentation as they elute from the LC system. This can e.g. be used in a discovery phase to be able to detect as many proteins as possible. If a targeted approach is desired, the acquisition mode is referred to as data-independent acquisition (DIA). A DIA approach is based on a pre-selection of the peptides that are attached as an inclusion list in the instrument method. A targeted approach will be more sensitive, especially for low-abundant proteins [168].

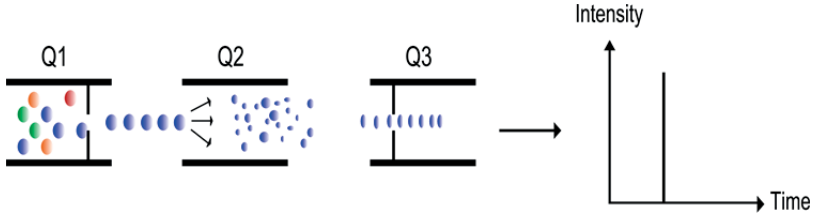

Figure 7. Selected reaction monitoring (SRM) in a triple quadrupole.

The most commonly used method for absolute targeted MS is multiple or selected reaction monitoring (MRM or SRM) in combination with stable isotope labeled peptides. Triple quadrupole (QQQ) is the instrument used in such analyses. The digested sample is introduced by a nano-LC system to the first quadrupole (Q1) that does a selection based on the m/z of a precursor from an inclusion list. The precursor is fragmented in the second quadrupole (Q2) and finally one or several of the fragment ions are detected in the last quadrupole (Q3) [171], Figure 7. Each such precursor-fragment pair is called transitions. SRM can be multiplex up to 1,000 transitions per run if a time-schedule SRM method is used. Time-scheduled SRM means that the inclusion list also contains a retention time window. This helps focusing the instrument on correct targets which will save instrument time for other transitions to be measured. To build a SRM method is usually time consuming due to the fact that transitions have to be selected. A good transition include a good peptide and that peptide should give a nice fragmentation pattern. To do this, selection can be based on in-house data of the analytes in form e.g. shotgun runs. There are also growing resources of publically available MS databases that can be helpful to find a good peptide for the quantification. Such resources can be the earlier mentioned GPMDB [162], PaxDb [163], PeptideAtlas [164], PRIDE [165] or ProteomeXchange [166]. Such resources can also be specialized databases for already established SRM methods such as SRMAtlas [199], MRMaid [200], MRMer [201] and MaRiMba

[202]. SRM is a widely used method but the time to set up a method can be an obstacle even for an experienced laboratory.

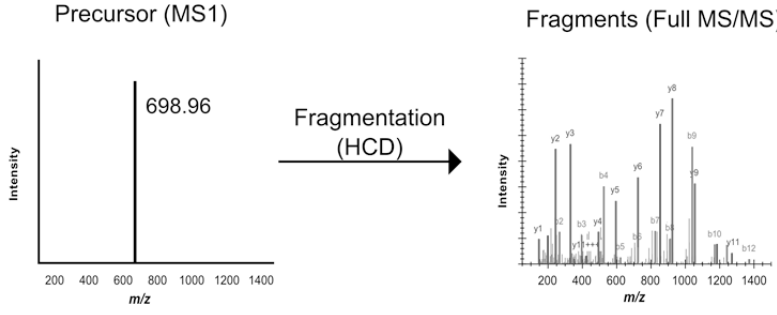

Figure 8. A PRM set up has similarities to a SRM analysis but also advantages. These are higher resolution and a full MS/MS acquisition. The full MS/MS means that the fragment selection can be performed post- acquisition.

In this thesis a recently presented approach for targeted quantification was used for quantification of feline insulin-like growth factors and two of their binding proteins (**Paper VI**). The method has been known as parallel reaction monitoring (PRM) and instead of a triple quadrupole instrument, a Q Exactive instrument is used for the analysis (a QTOF could also be used) [203-205]. The selection is also done in a quadrupole, but the fragmentation is done in a HCD cell. Finally the fragments are detected in an Orbitarp analyzer. The Q Exactive instrument gives the advantage, that instead of a few fragments detected, a full MS/MS spectrum is collected, see Figure 8. Which allows selection of fragments for the quantification to be done post-acquisition. This will save time during method development, because no specific transition has to be selected. There is, however, still need of a thorough selection of good peptides (precursor) which have to be established by in-house runs or the use of public database searches. Method optimization is also needed for several parameters such as automatic gain control (AGC), maximum inclusion time, resolution, isolation window and normalized collision energy (NCE).

A crucial part of the work with all types of mass spectrometry is the data analysis and targeted absolute quantitative MS is no exception. Rather, it is even more important in order to get a final result that is satisfactory. In this thesis an open-source software called Skyline [206] was used for the PRM data evaluation. It has been shown that PRM is capable to give limit of quantification in the low attomol range [207].

Targeted Protein Quantification – Mass Spectrometry versus ELISA

Targeted protein quantification has for a long time been done by ELISA, which still is the "golden standard" much due to a very good sensitivity. An obvious drawback with ELISA is the very long time for development of a new assay. A really sensitive ELISA is dependent on two very well validated antibodies. Another drawback is the lack of multiplexing abilities. SRM has been shown to routinely reach limits of detection at low attomol or ng/ml level [208, 209]. PRM has been shown to have limit of quantification in the low attomol range [207]. This is still not enough if e.g. interleukins in the low pg/ml level is to be measured [7]. However, there is a very good potential that, in a near future, this can be achieved due to the fast instrument progress that have been seen the last decade [210]. This means that MS targeted quantification is closing in on ELISA. Leigh Anderson points out, flow through as a very important aspect regarding MS based quantification compared to ELISA [210]. Today it is laborious and time consuming to analyze many samples with MS even if that analysis can give multiple answers for a large number of analytes. In ELISA it is the opposite. It is fairly doable to measure thousands of samples, but without multiplexing of analytes. To analyze 10 proteins in 10,000 samples with MS are simply not done due to cost and time. Anderson claims that he has a solution for the slow sample processing in MS. That would be automation and the use of a method he was one of the inventors of, SISCAPA. The application Anderson present, also circumvents the use of chromatography, simply because the sample complexity is reduced that much, there is no need. This shows one possible way, how to get the MS analysis into the clinical lab. There will for sure be more similar applications presented the coming years.

A well-defined biomarker discovery project based on MS could be outlined as follows. It would involve a discovery phase with shotgun runs and in this discovery phase it would be beneficial to include low-cost methods such as relative and label-free quantification to be able to reduce the number of candidates. Next step would be a hypothesis driven targeted quantification with some kind of isotopically labeled standards used as internal reference, preferably PRM or SRM. For an orthogonal verification of the findings, protein arrays, ELISA or Western blot, could be used. Even with a well-defined study set up, the path from idea to a clinical useful biomarker is long [211, 212]. This work has to be faster and more streamlined. With the enormous method developments that have been done the last two decades in proteomics there is a very good opportunity to make that change. It has to be done in an international community to be able to set up routines and validated workflows that can be used and evaluated by the whole community. Another crucial step is the infrastructure for data handling and data comparison. This will be the greatest challenge for the next two decades.

4. Affinity Enrichment and Depletion Combined with Mass Spectrometry

Combining affinity and mass spectrometry based proteomics can be beneficial in many ways. Verification of proteomics results is crucial. If it is possible to reproduce the results with both mass spectrometry and affinity based methods, the evidence for a correct observation is much more likely than produced by only one of them. Such verifications are truly orthogonal, which means that they are based on a completely different principle of measurements. Running a Western blot is not an orthogonal verification of an antibody array experiment, but it is orthogonal to a MS analysis. To reproduce the result with the same type of analysis is also valuable but does not give that much extra evidence that the findings are true discoveries.

Another way to take advantage of these two fields of proteomics is to combine them to get a better coverage or to get much more distinct results of the analysis. The dynamic range of protein concentrations is large in body fluids, up to 12 orders of magnitude [7]. This is a challenge both for mass spectrometry and affinity proteomics. LC-MS/MS analysis of a digested blood plasma sample will not cover the whole proteome. To produce an antibody array that could do so would be extremely expensive and a specificity of such a large setup can really be questioned.

Multiple Fractionation in Front of Mass Spectrometry

To be able to increase the coverage in mass spectrometry, there is a huge number of applications that have been tested and published and some of them will be summarized here. Many of them combine immunoaffinity and mass spectrometry, to take advantage of their strengths. Another way to deal with this issue is to combine two different LC methods in front of MS referred to as 2D-LC or a combination of SDS-PAGE with an ordinary reversed phase-high-performance liquid chromatography (RP-HPLC). A widely used technique is 2D gel and the combination with MS has made it possible to more accurately determine which protein there is in a certain spot. 2D gels were used long before MS was the analytical application of choice for proteins and is still a popular application [213]. For large scale proteomics studies the combination SDS-PAGE or 2D-LC is more common. In an SDS-

PAGE set up the sample is separated on the gel and the gel lanes are divided in smaller pieces and an in-gel digestion is done and finally analyzed in an ordinary reversed-phase LC-MS/MS. In a 2D-LC set up the first LC separation is done off-line and the collected fractions are then analyzed in an ordinary reversed-phase LC-MS/MS. In one study a comparison between off-line RP-HPLC, strong cation exchange HPLC (SCX-HPLC) and SDS-PAGE followed by an ordinary on-line RP-HPLC was done [214]. The conclusion was that there are advantages and limitations for all tested combinations and that the easiest and most cost effective was the gel based approach. This statement holds for much of the proteomics applications, that there is no universal application that can solve all proteomic challenges. The methods have to be selected according to a specific study and according to the equipment available and funding.

Immunoenrichment and High Abundant Protein Depletion

The methods mentioned above are not protein specific in any way and the separation is done in a more general fashion. This can be an advantage in e.g. a discovery study when as many proteins as possible is the aim. If one would want to have a selection of a specific protein or a sub group of proteins, an affinity approach could be one way to handle this. There is a large number of methods where this is applied. One has to be aware that the selection will not be any better than the specificity of the affinity binder that is used. I will come back to this issue but there is an ongoing debate about the quality of the enormous amount of antibodies and other types of affinity binders that are available for the research community [25-27, 66, 78-81].

Using affinity reagents to enrich or deplete samples in front of a mass spectrometry analysis can be conducted in several different ways. It can be in form of a single protein enrichment or depletion. It can be also be an enrichment/depletion of several analytes. There are many different methods and there are also a lot of suppliers of different assay kits which makes it a challenge to find an assay that really works effectively. The cost of a kit can be significant, but that does not automatically mean that it will produce a desired result. The enrichment may be conducted either on protein or peptide level. The enrichment strategy determines which characteristics an antibody should have. If the enrichment takes place at the protein level the antibody should be able to bind epitopes of the intact protein and should therefore preferably be produced against a full-length protein.

Immunoprecipitation as mentioned above is one way of doing enrichment. In this method, an antibody is immobilized onto a solid support. In a validated assay it specifically binds the targeted protein, thus concentrating

the protein [215]. This fraction can then be analyzed with LC-MS/MS either directly or after an SDS-PAGE run as was done in **Paper V**. Nelson and coworkers 1995 presented a setup, which they referred to as mass spectrometric immunoassay (MSIA) where they did IP followed by MALDI-MS identification [216]. If the enrichment instead is done on peptide level there has to be anti-peptide antibodies produced. Anderson et al. have presented a concept called stable isotope standards and capture by anti-peptide antibodies (SISCAPA) [217]. Enrichment has also successfully been used in the exploration of post-translational modifications. More about enrichment strategies both for proteins, peptides and PTMs can be found in Weiss et al. and Zhao & Jensen [23, 218].

A device of interest when talking about enrichment is the integrated selective enrichment target (ISET) which first was presented by Ekström and coworkers in 2004 [219]. With this device all sample preparation including enrichment, digestion and washing can be done in the same place minimizing the losses. Finally the analysis with MALDI can be done directly on the device. The format of ISET makes it possible to integrate it into a robotic handling system [219, 220].

There are also more global approaches for protein or peptide enrichment. For intact proteins hexapeptide combinatorial ligand libraries can be used and theoretically they should be able to bind a complete proteome [221]. This is a bead based system where the ligands are immobilized on beads and the sample that is to be enriched is added. The principle is that abundant proteins rapidly saturate their binding sites and the rest is washed away. Low abundant proteins are instead concentrated on the binding sites and the overall dynamic range of concentration is reduced.

For peptide enrichment there are two similar concepts that have been presented in the last years. Those are Global Proteome Survey (GPS) [222, 223] and Triple X Proteomics (TXP) [224, 225]. These two methods are based on binders that can capture short peptide motifs (3-6 amino acids). The motifs are selected using bioinformatics tools and they are shared by a few up to hundreds of proteins. In the GPS method synthetic peptides are generated and the context-independent motif specific (CIMS) antibodies are expressed using a large human recombinant antibody phage display library. In TXP the antibodies are expressed as polyclonal antibodies in rabbit. The idea to use motifs recognized by more than one protein is that the number of capture reagents needed to cover a proteome is reduced and it will also create a species independent enrichment method. When the antibodies are produced they can be used for capturing of peptides from a digested sample. A big drawback is that there is mainly one peptide for each protein that are detected which puts demands on the post-analysis to lower number of false positives. The concept is promising and it will be interesting to see how these techniques evolve.

There are a number of different capture agents used in the high abundant protein depletion kits found on the market. Most dominantly are antibodies and IgG are the most used, there are also hen antibodies (IgY), affibodies, Protein A and G, Cibacron blue and Blue Dye [226-230]. The antibody based should be more specific and more effective than the ones based on an affinity matrix for target removal. The antibody based assays are reported to deplete from 95 up to 99% of the targeted proteins, while the other reports depletion between 80-95% for human samples [226]. Albumin is the most dominant protein in human blood plasma and account for about 50% of the total content of plasma proteins [231]. There is a risk that proteins that are not supposed to be depleted anyway are depleted due to nonspecific binding or bound to depleted proteins [232, 233]. This means that the depletion not necessarily has to be that reproducible. High abundant protein depletion is not a particularly good tool for quantitative proteomics, then a specific selection of one or a number of proteins will be targeted and that is a better choice. The capture agents are immobilized on agarose beads or similar material. The assays are based on spin columns, gravity columns or LC-columns depending on the manufacturer and the throughput and amount of sample that have to be handled. The number of proteins that can be depleted with different assays varies from one (albumin) up to 20 of the highest abundant proteins in human blood plasma [234]. Even if there are restrictions of the effectiveness and reproducibility by using high abundant protein depletion kits the complexity of the sample will decrease. This will be helpful in detection of more proteins and specific proteins that cannot be detecting by e.g. multiple fractionation. Assays are developed for human blood plasma but they have been used for human CSF [235, 236] and there are also publications where assays have been used for depletion in other species [237, 238]. For the model animal rat and mouse there are specific assays developed [239, 240].

In this thesis high abundant protein depletion was used in combination with SDS-PAGE and MS analysis in **Paper VII** as an effort to improve the coverage of dog cerebrospinal fluid (CSF) proteome. High abundant protein depletion in combination with SDS-PAGE [235] and SCX-HPLC [236] has been successful in the mapping of the human CSF proteome. The aim of the dog study was to investigate if it was possible to use depletion assays developed for human or rat blood plasma sample on dog CSF. The study showed that it is possible to use the tested protein depletion kits for dog CSF sample. All four tested kits could deplete dog proteins, but the efficiency varied. The best performing kit showed a very good improvement for proteins with similar molecular weight as albumin and a total of 75 more proteins were found in that fraction using the depletion kit. This clearly shows that there is a value to use depletion to uncover more of different proteomes.

5. Contributions to the Field

When the HPA project started in 2003, dot blot was set up as the standard method for validation of the specificity of the produced antibodies. This was not an optimal way to test the specificity. Therefore, in order to obtain a more reliable method, an antigen microarray was introduced. Figure 9 shows an antigen array consisting of 384 antigens compared with the dot blot set up. The development of the protein microarray accelerated during 2004. After a thorough discussion about the robustness of the antigen microarray as a validation tool it was implemented. Antigen microarrays were the foundation of my work at KTH and a continuous development of both the equipment and the assay itself was performed.

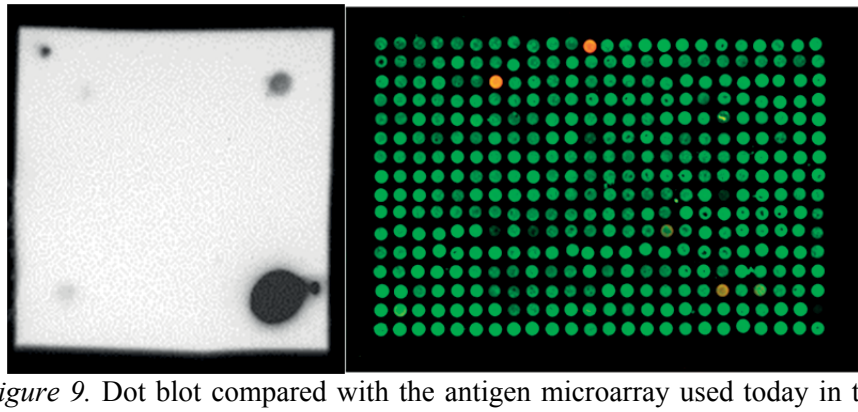

Figure 9. Dot blot compared with the antigen microarray used today in the HPA project.

In 2013 I was given a fantastic opportunity to broaden my knowledge within the proteomics field. I got the chance to learn about mass spectrometry based proteomics. I thought that some of the knowledge from my chemical engineer education could be useful. The answer is a very weak yes, a really tiny part was useful. The rest was totally new. It has been a challenge to get into mass spectrometry based proteomics in about two and a half year. It has been very exciting years because I have learnt how powerful and useful mass spectrometry can be in the proteomic research. I have only started to scratch on the surface of what is possible to do with mass spectrometry. However, I have been able to show that mass spectrometry can be a useful tool in animal proteomics both as a discovery and a quantification tool. The proteomics

field is a fast growth with new achievements reported every day. It has been a privilege to be a small part of it both from mass spectrometry and affinity based proteomics.

Validation of Affinity Reagents – Using Antigen Microarrays or Mass Spectrometry

In **Paper I-III,** three approaches used to test the specificity of the HPA produced antibodies are presented. In **Paper I** the antigen microarray was used and that was the method used in the HPA project. The antigen microarray has the advantage that it can theoretically be scaled up to thousands of proteins immobilized and the printed microarrays were stable for long time. The CD based platform in **Paper II** gave a more differentiated picture of the specificity of the antibody. The bead based approach in **Paper III** was proven to be very flexible in the sample handling and combination of antigen. The use of PrEST based microarrays for antibody validation proved to be a very good tool and the approved antibodies were functional to a large extent in both immunohistochemistry and in immunofluorescence. The choice of planar arrays as the preferred analysis tool in the HPA project was based on two parameters. Cost and that the workflow was already established when the two other applications were set up. The Gyros system should not have been applicable in the high-throughput of the HPA project. The bead based system could have worked and should most likely have been easier to set up in a more streamlined and automated fashion than the planar arrays. Cost could have been higher due to the cost of beads. Planar antigen microarrays are working horses in the HPA project and have been a valuable validation tool of the produced antibodies. To increase the flow through and give a more accurate estimate of the specificity, the microarrays were later in the project scaled up to 21 sub arrays, consisting of 384 antigens, in one slide. The use of protein arrays for antibody specificity validations have also been reported by other groups as an effective approach [241-245].

Paper IV is about the usefulness of the in HPA developed antigen microarray, in a broader perspective. The antibody validation presented in this article was a part of the SH2-consortium [38]. This consortium was established to explore the prospects of a systematic development of high-quality antibodies to human antibodies. Most SH2 domain-containing proteins are recruited to phosphorylated tyrosine sites and are thereby deeply involved in cell signaling [246-248]. To have well validated renewable affinity binders for these proteins would be highly valuable for the research community.

In the study presented, the planar antigen microarray was further developed to be used as a validation tool for other kinds of affinity reagents, produced externally. In the study, three different kinds of affinity reagents were

included; mono-specific HPA antibodies, mouse monoclonal antibodies and recombinant single-chain variable fragment (scFv) that were produced by six different academic laboratories. A protein microarray was constructed including, 105 SH2-antigens and 301 non-SH2 proteins. Twenty of the SH2-antigens were produced by the Structural Genomics Consortium (SGC) [249] at Karolinska Institute. The functionality of the PrESTs was known, while the behavior of the SGC proteins was unknown. When the affinity reagents were tested, it was shown that the SGC proteins could be printed in the same way as the PrESTs and their functionality was equal to the PrESTs.

All in all, 398 affinity reagents were tested, of those about 50% had a binding profile with highest signal for the targeted antigen. Only 10% had no or low signal for the intended antigen while reactivity against other antigen could be detected. This study clearly showed that it was possible to use the HPA antigen microarray for testing of other binders than in-house produced. It was also shown that it was possible to include antigens produced in a different workflow. Further validation of the antibodies with e.g. immunoprecipitation followed by MS could have revealed if the binder that could distinguish between ABL1 and ABL2 was as specific as indicated.

In **Paper V**, 13 antibodies that were developed to target the steroid hormone receptor, Estrogen receptor β, were validated using three different assays. A first screening was done using tissue microarray which resulted in that only three of the antibodies (PPZ0506, 14C8 and PPG5/10) were further analyzed, using Western blot and IP-MS. A clinically used antibody targeting Estrogen receptor α was also tested as a positive control. I was responsible for the mass spectrometry analysis of the antibodies. The mass spectrometry analysis resulted in the fact, that only PPZ0506 confidently bound Estrogen receptor β. For the antibody 14C8, Estrogen receptor β was only found in a search against a smaller set of ER related proteins i.e. a more targeted search. Instead 14C8 seems to target a protein from the POU transcription family. The antibody PPG5/10 is an antibody that has been used in numerous publications. In our hands it showed a very unspecific behavior when tested in IP-MS. No detection of the correct protein could be recorded, but over 40 other proteins were detected in the two runs that were conducted. Eleven of these antibodies were commercial antibodies, many of them used in different ERβ publications. This, once again, shows how difficult it can be to find really good antibodies on the market. As have been said before, there are many examples of studies that have shown large unspecific behavior of commercial antibodies [76, 77] or studies that have been spoiled by bad antibodies [74]. In this study we showed that a combination of IHC, WB and MS was a powerful tool to investigate the specificity of commercial antibodies. The methods gave consistent results of the tested antibodies. A bad performing WB resulted in an as bad MS analysis using the antibody.

Mass Spectrometry as a Tool in Animal Proteomics – Supported by Affinity-based Applications

There is a growing interest in animal proteomics, not only from veterinary researchers. This was reflected at the last HUPO conference in Madrid, Spain, where Non-Human and Food Proteomics for the first time had an own session. Animal proteomics will have a big impact also for humans since there is a lot of model organisms used in the study of human diseases. Some examples are dogs with similar diseases as humans such as neurological related diseases and Alzheimer´s disease [250-252]. Then there is the well-known use of mouse, rat, zebrafish as model animals for studies of different human diseases [253-255]. Another large field is farm animal proteomics [256-258]. A better understanding of the model animals and the animals that we have a close relation to would benefit both areas. Proteomic studies can uncover both differences as well as similarities. The explosion of assays and reagents in the affinity proteomics field that have been seen the last decade for humans, have not been witnessed in animal proteomics. When affinity assays are used within the animal proteomics field, it is common to use human developed assays adapted for animal use. This can be successful but there are huge challenges to handle in human affinity proteomics and it will not be any easier trying to adapt for animal use. Mass spectrometry is already used in animal research and could be used even more. A combination with affinity proteomics will be useful in many ways. It can be used for large scale discovery studies, where both applications are used to get a better coverage. It can also be used as a validation tool, to verify that the used assay actually gives a correct answer. The usefulness of mass spectrometry in animal proteomics was investigated in **Paper VI** and **VII**.

A targeted absolute quantification mass spectrometry-based approach was investigated in **Paper VI**. Targeted quantification of proteins has been done with ELISA for a long time and is regarded as the “golden standard”. This is also true for veterinary practise but the problem is that there are few specific ELISA assay for animals. This can create problems with unspecific binding, with measurements giving too high or too low values. **Paper VI** demonstrates the applicability of a parallel reaction monitoring (PRM) method using QPrESTs for quantification of four feline serum proteins. PRM uses a high-resolution Orbitrap MS instrument (Q Exactive Plus) for the targeted analysis. We could show excellent quantitative sensitivity (in low femtomol range) and selectivity for all four growth hormones. A good correlation with the golden standard ELISA method was proven. Further, it was shown that the QPrEST, developed to mimic human proteins, can also be used in animal proteomics as internal standard. The study shows that a targeted quantitative mass spectrometry based method can be a valuable tool in the veterinary research.

In **Paper VII,** the aim was to investigate if high abundant protein depletion columns developed for human/rat sample could be useful in mapping the dog cerebrospinal fluid (CSF) proteome in combination with mass spectrometry analysis. CSF was chosen because there is a long tradition of analysis of this body fluid in the context of neurological diseases in the department. Dogs can have similar neurological diseases and are therefore used as model animals in e.g. Alzheimer´s disease [252]. It would therefore be beneficial to have a better knowledge about the dog CSF proteome. Today only 811 proteins of the estimated 25,485 proteins in the dog proteome are reviewed in UniProt. The rest of them are, to a large extent, homology estimations. There is a lot more to learn about the dog proteome! In the study, a total of 983 proteins were detected, of which 801 proteins were marked as uncharacterized in UniProt. It was shown that all of the tested columns were able to deplete proteins. There were large differences between the columns. The best performing column, regarding the number of detected proteins, repeatability and ability to deplete the targeted proteins, was antibody-based. The largest effect with this column was seen for proteins of similar size as albumin. It was also shown that fractionation of the sample was a great way to increase the coverage, which has been shown in human studies too [235, 236]. The challenge with extensive fractionation is that the time for mass spectrometry analysis will be very long. If the aim is to cover as much as possible of the proteome it is manageable. However, if the aim is to have a large scale study with many different samples it will be more or less impossible. It was shown that the use of depletion columns, constructed to deplete human or rat proteins, increase the number of detected proteins in dog CSF samples.

Conclusions and Future Perspectives

We are in a dynamic area of proteomics and there will be many exciting discoveries in the coming decades. There have to be more large scale collaborations because it is a very complex research field with a lot of challenges. There have been a tremendous increase of knowledge of both plants and animals, including humans in the last decades, but there is still an ocean of knowledge to be harvested. A closer connection of affinity and mass spectrometry based proteomics will strengthen both. There is another challenge that has to be handled and that is how all the gathered information should be stored and shared to benefit the most. The infrastructure of handling the proteomics data is lagging behind and has to be improved a lot to speed up the work of improving the diagnostic and treatment of diseases that affect humans, other animals and plants. This is valid for both the success stories as well as the less successful projects. If there is better reports of failures it is

more likely that another research group can use that information to re-estimate the problem and hopefully find a solution.

Another issue with the proteomics field is the validation of affinity reagents or the lack of that. It is good that this topic is more debated and that researchers are more aware of the shortcomings that could appear using affinity reagents. A better awareness of the responsibility among researchers to validate the finds and the reagents that are used, is also requested. Today there are no general agreements about how to validate affinity reagents and this leads to substantial waste of valuable resources and restricts the growing market of affinity based diagnostics. Therefore, better international agreements on how affinity reagents should be tested to be regarded as functional reagents are needed. One of the most important issues is the specificity of affinity reagents. There has to be agreement in what kind of application specific validation that is needed, because different assays need different validation of the antibodies used. It would be too expensive to have the same validation for all antibodies that are available but there has to be a basic level. Then the suppliers should specify for which assays their affinity reagents are recommended and have appropriate validation for that.

In this thesis, it is shown that the antigen microarray platform that was established within the HPA project was a very good tool to determine the specificity of different affinity binders. The antigen microarray could be used in future collaborations as a tool to produce specific affinity binders against different targets. In the presented papers, it has been shown that the assay is not only usable with the antigens and affinity binders that are produced in the HPA project but also with reagents produced by other research groups around the world. In addition to the antigen microarray it was shown that mass spectrometry is a very good method for validation of antibodies. In a study including IHC, WB and mass spectrometry there were good correlation of the results concerning the specificity for the tested antibodies for the different methods.

There is a need of methods recognized as reliable by the community, for validation of antibodies. The methods have to be able to handle different kinds of affinity reagents produced at different laboratories. The methods should be so easy to use and as robust that there is no difference by whom and where the assay is carried out. There is still much work needed before such generally accepted and standardized methods are established. There are promising indications showing that we are heading in that direction, e.g. by researchers publishing articles where they demand better and more standardized methods for affinity reagent validation. Portals like antibodypedia.org and pabmabs.com can be valuable to put pressure on manufacturer to withdraw antibodies with bad performance. To get there, such resources have to be better known by the research community and the information published has to be critically evaluated. If such portals would be curated by a non-

profitable organization with limited connections both to the academic world as well as the commercial market it would strengthen their credibility.

More efforts have to be invested in using orthogonal verifications of biological findings. A verification of an MS study with Western blot is very good but it would be even better if it was performed with some kind of protein microarray set up. There are examples of using mass spectrometry to validate bead-based antibody array results [259, 260]. So far, sensitivity has been an issue resulting in large amount of beads used for the MS verification. Hopefully these kinds of attempts will be more common and these kinds of problems will be possible to overcome.

In this thesis, I have shown that mass spectrometry can be very helpful to get a better understanding also in animal proteomics. The dog CSF study shows that with a combination of different analytical methods it is possible to uncover many of the proteins that can be found in dog CSF. Dog is a model animal for human diseases and with mass spectrometry it is possible to do multi species analysis. I also showed a targeted quantitative mass spectrometry based method with good sensitivity.

ELISA has historically provided better sensitivity than MS-based methods. Now, however, MS is getting closer and in combination with a much better specificity, there is much to gain using MS in animal proteomics. If the protein to be quantified is very homologous between different species, then there would be a possibility to set up a MS-based targeted assay that could be used for many different species. So far, mass spectrometry is a very complex and knowledge based application that is challenging to implement in a clinical laboratory and especially in a veterinary laboratory. As a research tool it can already give valuable new information about various animal proteins and their function and similarities to human ditto. SRM or PRM methods can be used in a clinical context and there are studies showing the possibilities with MS-based methods [261]. It will be more common with MS-based methods in the clinical laboratory but still the immunoassay based perspective is the most common [262, 263]. For veterinary practice it would be beneficial to have mass spectrometry-based assays and it would be easier to implement new analytes due to e.g. shorter time for method development. Still it will take many years before mass spectrometry will be used more widely in veterinary practise due to the cost and complexity of the method.

Mass spectrometry harbors a great potential to be an even more invaluable analytical tool both in the research laboratories and in clinical laboratories. It will for sure be beneficial to combined mass spectrometry and affinity proteomics.

Summary in Swedish

Proteomik är ett forskningsfält inriktat på studier av proteiner. År 2001 presenterades den första fullständiga kartläggningen av människans genom. Det vill säga människokroppens lagringssystem av information om vårt genetiska arv. Genom att knäcka denna kod kunde forskare börja kartlägga varför vi ser ut som vi gör, hur vi fungerar m.m. Framförallt fick vi för första gången tillgång till koden för alla de proteiner som kroppen är uppbyggd av. Förhoppningarna var högt ställda i början för att lösa den ena gåtan efter den andra med hjälp av den genetiska informationen som finns lagrad i våra kroppar. Tyvärr visade det sig att kroppen är lite mer komplex än bara den information som finns lagrad i våra gener. Den genetiska sekvensen ger inte all information om ett funktionellt protein. Under processen från översättning av gensekvensen till dess att proteinet är fullt funktionellt sker många olika modifieringar. Detta kan vara att gensekvensen klipps på olika sätt, vidare är proteiner veckade i en 3-D struktur som ger olika möjligheter för proteinet. Slutligen finns det många olika molekyler som kopplas på proteinet under dess ”mognadsprocess” för att ge proteinet särskilda egenskaper. Naturen är en fantastisk byggmästare som har haft miljoner av år på sig att förfina de processer som gör att vi fungerar som människor. För att lösa detta gigantiska pussel behövs mycket tid, mycket tänkande och framförallt metoder att kunna analysera det som händer i våra kroppar.

Sedan 2001 har vi med hjälp av olika proteomikmetoder lärt oss gigantiskt mycket mer om människokroppen och även om andra organismer, men vi är fortfarande på en basal nivå av förståelsen av det naturen skapat. Vi har insett att det inte alltid är helt lätt att veta vad ett protein gör, var det finns och i vilken mängd. Dessutom har vi fått en mycket mer mångfasetterad bild av hur proteiner modifieras efter att det har uttryckts med hjälp av den genetiska koden i våra celler. Dessa modifieringar har visat sig vara oerhört viktiga för hur proteiner signalerar, uttrycks och fungerar i största allmänhet. Den snabba utvecklingen av sekvenseringsmetoder för både DNA och RNA har resulterat i allt fler fullständigt kartlagda genetiska sekvenser, inte bara för människor, utan för snart sagt alla levande organismer. Det har gett en fantastisk möjlighet att storskaligt kartlägga proteinsammansättningen i alla de kartlagda organismerna på ett sätt som inte tidigare varit möjligt.

Masspektrometri är en mycket användbar metod inom proteinforskningen. En masspektrometer är en detektor som kan användas för att bestämma massan och laddningen på de molekyler man vill analysera. Med hjälp av

dessa och kunskapen om hur proteinsekvenserna är uppbyggda av olika aminosyror kan man sedan ta reda på vilka proteiner som finns i det analyserade provet. Ofta är det dock frågan om väldigt komplexa prover med väldigt stort spann mellan de proteiner som förekommer i störst mängd och de som är minst förekommande. För att bättre kunna analysera vilka proteiner som finns i ett prov använder man sig av vätskekromatografisk separation i kombination med masspektrometri. Vätskekromatografi utförs med hjälp av en kolonn innehållande en stationärfas. Till denna kolonn är kopplat en eller flera pumpar som levererar det som kallas mobilfas. Det är sedan egenskaperna på den stationära och mobila fasen som avgör den slutliga separationen av komponenterna i provet. Ett vanligt sätt inom masspektrometri är att separera med avseende på hydrofobicitet. Det är dock inte alltid det räcker med vätskekromatografisk separation och då kan olika molekyler med förmåga att specifikt fånga upp ett protein, här kallade bindare, användas för att på så sätt förbättra den slutliga masspektrometrianalysen. Det kan vara fråga om att framställa naturens egen lösning på detta som kallas för antikropp eller så kan det vara på laboratoriet modifierade varianter av dessa antikroppar eller andra lösningar för att fånga upp och rena fram ett protein från en komplex blandning av många proteiner. Forskningsområdet som baseras på antikroppar och liknande bindare brukar benämnas affinitetsbaserad proteomik. Det finns väldigt många sätt att producera dessa ”fångstredskap”. För att de skall fungera optimalt måste man vara säker på att de binder upp rätt protein och att de inte samtidigt även binder upp andra proteiner. I kroppen finns utmärkta regleringssystem för att se till att antikropparna hittar rätt och att de gör vad de skall göra. I kroppen fungerar antikroppar som ett sorts flaggsystem för att visa främmande substanser eller organismer för kroppens immunförsvar så att dessa substanser kan rensas bort från kroppen.

I den här avhandlingen har både masspektrometri och antikroppsbaserade metoder använts, såväl den ena metoden var för sig som de båda i kombination. I **artiklarna I-V** är utgångspunkten att försöka verifiera att bindare som har producerats av forskare på laboratorier verkligen binder till det protein som de var konstruerade för att binda. Det görs dels med hjälp av något som kallas för proteinmikroarrayer (**artiklarna I-IV**) men även med hjälp av masspektrometri (**artikel V)**. I de proteinmikroarrayer som används i denna avhandling binds det protein som användes för att producera bindaren upp på en fast yta (mikroskopglas eller kula). För att kunna fastställa att en bindare är specifik i sin bindning behövs många proteiner att testa den. I denna avhandling användes upp till 384 (målproteinet plus 383 andra) olika proteiner för detta ändamål. När dessa proteinmikroarrayer är producerade så kan analysen av bindare börja. Med en specifik färginmärkt molekyl kan man se om man har lyckats fånga upp bindaren med något av proteinerna. Om bindaren är specifik så får man i detta fall en röd färgmarkering för det protein den är tillverkad mot. Många gånger fanns dock en svagare infärgning även för andra proteiner, se Figur 9. För att lättare kunna lokalisera var de olika prote-

inerna fanns på mikroarrayen användes en annan molekyl som var specifik för en sekvens som fanns hos samtliga proteiner (grön färg). Tre olika tekniska plattformar användes för att sätta upp metoden för att verifiera att de antikroppar som producerades i ”Human Protein Atlas (HPA)” projektet (**artikel I-III**). Det visade sig att de gav väldigt samstämmiga svar för de analyserade antikropparna. I **artikel IV** visades att den proteinmikroarray som användes för validering av alla antikroppar i HPA projektet även kunde användas för validering av bindare producerade av andra forskargrupper. I **artikel V** visas att masspektrometri kan vara ett mycket användbart verktyg för att ta reda på om en antikropp är specifik eller inte. I den studien gjordes en masspektrometrisk analys efter att antikroppen hade fått möjlighet att binda upp det protein den var designad att binda, målproteinet. Endast en av de tre antikroppar som testades med hjälp av masspektrometri för ett specifikt protein band till det proteinet. Vidare kunde konstateras att de två övriga hade bindningar till andra proteiner varav en av antikropparna band till väldigt många andra proteiner. De resultaten överensstämde väl med de andra analyserna för att kontrollera dessa antikroppars specificitet. Styrkan med masspektrometri är att analysen också gav svar på vad för andra proteiner det kunde vara frågan om då inte målproteinet kunde detekteras.

Att masspektrometri är ett mycket värdefullt verktyg inom proteinforskningen exemplifierades vidare i två studier som inriktades på prover från hund respektive katt. I hundstudien (**artikel VI**) analyserades cerebrospinalvätska (ryggmärgsvätska) för att försöka detektera så många proteiner som möjligt i denna vätska från hund. Här användes en kombination av kommersiella kit som utvecklats för att fånga upp specifika mänskliga proteiner i kombination med massspektrometri. Idén med att använda bindare här var att försöka ta bort så mycket som möjligt av de mest förekommande proteinerna för att på så sätt lättare kunna detektera de som förekommer lite mer sparsamt. Ingen hade tidigare utvärderat om metoderna kunde användas för beredning av hundprover. Vi kunde visa att det var möjligt att använda dessa kit för hund cerebrospinalvätska. Här lades även ett extra separationssteg in i form av en gelseparation. Enkelt beskrivet laddas provet med proteinerna på en i förhand gjuten gel. Därefter läggs en spänning över gelen och proteinerna i provet separeras med avseende på storlek i gelen. Totalt kunde 983 proteiner detekteras i detta prov med hund cerebrospinalvätska. I kattstudien (**artikel VI**) visas hur masspektrometri kan användas för att bestämma den faktiska koncentrationen av olika proteiner i ett prov. Principen bygger på att man tillsätter ett i laboratoriet framställt protein med en något annorlunda massa än det naturliga proteinet man vill studera. Eftersom det är framställt i en laboratoriemiljö kan man bestämma dess exakta koncentration. I vätskekromatografisteget separeras dessa båda på samma sätt och kommer därmed att introduceras och analyseras samtidigt i masspektrometern. Däremot gör det faktum att de har olika massor att de kan separeras i masspektrometern. Eftersom koncentrationen på det konstgjorda proteinet är känt så kan

man sedan med hjälp av det bestämma vilken koncentration det är på proteinet i provet. Några av fördelarna med en masspektrometrimetod är att det går att mäta koncentrationen på flera proteiner i samma experiment och att den inte är djurspecifik utan relativt enkelt kan anpassas att mäta på andra djurarter om det inte är möjligt redan från början.

Jag har i denna avhandling i en studie visat att masspektrometri är ett kraftfullt verktyg för att bedöma specificiteten på antikroppar och i två studier dess värde i proteinanalyser för hund och katt. Vidare har jag visat att en kombination av masspektrometri och molekyler som kan fånga upp proteiner kan vara en mycket användbar kombination. Båda metoderna har fördelar och nackdelar. Med en större medvetenhet om dessa och ett bättre utnyttjande av de båda applikationerna kommer vi att uppleva många nya och spännande upptäckter om hur proteiner fungerar och verkar i naturen.

Acknowledgements

Jag vill börja med att säga att det varit en lång och krokig resa. Jag kommer inte att ha möjlighet att tacka alla underbara människor personligen eftersom det genom åren har passerat en stor mängd av väldigt smarta och trevliga människor.

Jag börjar med mina handledare genom åren:

Margareta, tack för alla diskussioner om både forskning och annat världsligt. Dina många gånger välfyllda kommentarer på manuskript har varit lärorika och berikande, men även en utmaning. Du har varit en superbra handledare som är omtänksam och stöttande.

Tack **Jonas** för att du gav mig möjligheten att ta upp mitt doktorerande och därmed möjligheten att få en inblick i den fantastiska MS världen.

Peter och **Jochen** som hjälpte mig att kryssa fram i array universum på KTH.

Sandra, Anna A, Cecilia W och Figge för diskussioner om Erβ, IHC och mycket mer som rör antikroppar som inte riktigt håller vad de lovar. **Emma och Bodil** för att ni har sådan glöd när det gäller forskning om katter och hundar. Jag har lärt mig massor. Ett stort tack även till alla andra medförfattare som har gjort det så intressant att komma till jobbet.

To all the room-mates that I have had during all these years. A special thanks to my present room-mates **Hilde and Neil**. It is nice be able to share both success and failure. Good luck, I am confident that you will do a very good work. Speciell tack även till **Kicki** och **Annelie W** som jag delade rum med så många år.

Katarina H och Torgny för mång och långa diskussioner om Pinpoint och Skyline. Jag skulle ha stampat på samma ställe ännu om det inte vore för er.

Sara för att du alltid tar dig tid att lyssna och är så omtänksam.

Anne for all the help to get me up and running with the Q Exactive.

All the fantastic fellow PhD students at the Department of Analytical chemistry: **Julia, Lena, Torgny, Ping, Mikael, Amelie, Anne-li, Sravani, Marcus, Hilde, Neil and Malin.** Thank you all other **colleagues** at the Department of Analytical chemistry for sharing your knowledge and guidance in moments when I have been confused.

Alla tidigare kollegor inom HPA projektet. Det är så många härliga människor som jag har mött under åren. Först och främst vill jag tacka **Ronald** min trogne vapendragare inom PrEST arrayerna och **Anna G** som kom in och styrde upp verksamheten. **Rebecca** for a very enriching collaboration. **Hanna T** som lotsade mig in i HPA, **Marica, Bahram, Anna S och Anja** i Molbio. **Kimi, Burcu, Anna H, Maja och Ulrika** i PAPP gruppen. **Johanna S**, **Pia A**, **Anna K**, **Cajsa**, **Erik M**, **Ulla W**, **Jenny O**, **Henrik**, **Mathias**, **Sophia**, **Caroline**, **Figge** och **Kenneth** och självkart alla andra som jobbat i HPA projektet eller på bioteknologi, plan 3. Jag har lärt mig så oerhört mycket under åren på KTH.

Till alla mina **vänner och släktingar**, nära och långt borta som jag har haft så mycket glädje med utanför forskningen. Det är skönt att veta att jag inte alltid behöver prata forskning utan det kan vara barn, jakt eller bara livet i största allmänhet.

Jag vill även passa på att tacka **Kalle** för att du tog dig tid att läsa min avhandling. Du kommer att kunna hänga med bättre men jag kan inte lova att du kommer att förstå allt på disputationen.

Tack även **Agneta** och **Hasse** för att ni har en sådan bedårande dotter och för att ni alltid är beredda att ställa upp om det behövs.

Mamma och **pappa** som alltid finns där när det behövs som mest. Tack för att ni alltid har ställt upp för mig. Det känns tryggt att ha er som jag kan luta mig emot.

Simon, min bror, och dina fantastiska barn. Vad skulle jag göra om jag inte kunde ringa dig när bilen krånglar eller när jag behöver prata av mig. Det är skönt att ha en som känner en nästa lika bra som man gör själv.

Min älskade **Katarina**, tack för att du är en sådan underbar människa. Du är det bästa som har hänt i mitt liv och så har vi ju fått två underbara killar. Jag älskar dig.

Viggo och **Joel,** nu är jag äntligen klar med boken. Nu lovar jag att jag har mer tid för er.

References

1. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid.* Nature, 1953. **4356**: p. 737-738.
2. Crick, F., *Central dogma of molecular biology.* Nature, 1970. **227**(5258): p. 561-3.
3. Cunningham, F., M.R. Amode, D. Barrell, K. Beal, K. Billis, S. Brent, D. Carvalho-Silva, P. Clapham, G. Coates, S. Fitzgerald, L. Gil, C.G. Giron, L. Gordon, T. Hourlier, S.E. Hunt, S.H. Janacek, N. Johnson, T. Juettemann, A.K. Kahari, S. Keenan, F.J. Martin, T. Maurel, W. McLaren, D.N. Murphy, R. Nag, B. Overduin, A. Parker, M. Patricio, E. Perry, M. Pignatelli, H.S. Riat, D. Sheppard, K. Taylor, A. Thormann, A. Vullo, S.P. Wilder, A. Zadissa, B.L. Aken, E. Birney, J. Harrow, R. Kinsella, M. Muffato, M. Ruffier, S.M. Searle, G. Spudich, S.J. Trevanion, A. Yates, D.R. Zerbino, and P. Flicek, *Ensembl 2015.* Nucleic Acids Res, 2015. **43**(Database issue): p. D662-9.
4. Gonzalez-Porta, M., A. Frankish, J. Rung, J. Harrow, and A. Brazma, *Transcriptome analysis of human tissues and cell lines reveals one dominant transcript per gene.* Genome Biol, 2013. **14**(7): p. R70.
5. Jensen, O.N., *Interpreting the protein language using proteomics.* Nat Rev Mol Cell Biol, 2006. **7**(6): p. 391-403.
6. Wilkins, M.R., C. Pasquali, R.D. Appel, K. Ou, O. Golaz, J.C. Sanchez, J.X. Yan, A.A. Gooley, G. Hughes, I. Humphery-Smith, K.L. Williams, and D.F. Hochstrasser, *From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis.* Biotechnology (N Y), 1996. **14**(1): p. 61-5.
7. Anderson, N.L. and N.G. Anderson, *The human plasma proteome: history, character, and diagnostic prospects.* Mol Cell Proteomics, 2002. **1**(11): p. 845-67.
8. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.
9. Tyers, M. and M. Mann, *From genomics to proteomics.* Nature, 2003. **422**(6928): p. 193-7.
10. Bjorling, E. and M. Uhlen, *Antibodypedia, a portal for sharing antibody and antigen validation data.* Mol Cell Proteomics, 2008. **7**(10): p. 2028-37.
11. Molgaard, S., M. Ulrichsen, S. Boggild, M.L. Holm, C. Vaegter, J. Nyengaard, and S. Glerup, *Immunofluorescent visualization of mouse interneuron subtypes.* F1000Res, 2014. **3**: p. 242.
12. Lander, E.S., L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J.P. Mesirov, C. Miranda, W. Morris,

J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J.C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R.H. Waterston, R.K. Wilson, L.W. Hillier, J.D. McPherson, M.A. Marra, E.R. Mardis, L.A. Fulton, A.T. Chinwalla, K.H. Pepin, W.R. Gish, S.L. Chissoe, M.C. Wendl, K.D. Delehaunty, T.L. Miner, A. Delehaunty, J.B. Kramer, L.L. Cook, R.S. Fulton, D.L. Johnson, P.J. Minx, S.W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R.A. Gibbs, D.M. Muzny, S.E. Scherer, J.B. Bouck, E.J. Sodergren, K.C. Worley, C.M. Rives, J.H. Gorrell, M.L. Metzker, S.L. Naylor, R.S. Kucherlapati, D.L. Nelson, G.M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D.R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H.M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R.W. Davis, N.A. Federspiel, A.P. Abola, M.J. Proctor, R.M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D.R. Cox, M.V. Olson, R. Kaul, N. Shimizu, K. Kawasaki, S. Minoshima, G.A. Evans, M. Athanasiou, R. Schultz, B.A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W.R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J.A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D.G. Brown, C.B. Burge, L. Cerutti, H.C. Chen, D. Church, M. Clamp, R.R. Copley, T. Doerks, S.R. Eddy, E.E. Eichler, T.S. Furey, J. Galagan, J.G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L.S. Johnson, T.A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W.J. Kent, P. Kitts, E.V. Koonin, I. Korf, D. Kulp, D. Lancet, T.M. Lowe, A. McLysaght, T. Mikkelsen, J.V. Moran, N. Mulder, V.J. Pollara, C.P. Ponting, G. Schuler, J. Schultz, G. Slater, A.F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y.I. Wolf, K.H. Wolfe, S.P. Yang, R.F. Yeh, F. Collins, M.S. Guyer, J. Peterson, A. Felsenfeld, K.A. Wetterstrand, A. Patrinos, M.J. Morgan, P. de Jong, J.J. Catanese, K. Osoegawa, H. Shizuya, S. Choi and Y.J. Chen, *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.

13. Venter, J.C., M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, J.D. Gocayne, P. Amanatides, R.M. Ballew, D.H. Huson, J.R. Wortman, Q. Zhang, C.D. Kodira, X.H. Zheng, L. Chen, M. Skupski, G. Subramanian, P.D. Thomas, J. Zhang, G.L. Gabor Miklos, C. Nelson, S. Broder, A.G. Clark, J. Nadeau, V.A. McKusick, N. Zinder, A.J. Levine, R.J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M.

Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A.E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T.J. Heiman, M.E. Higgins, R.R. Ji, Z. Ke, K.A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G.V. Merkulov, N. Milshina, H.M. Moore, A.K. Naik, V.A. Narayan, B. Neelam, D. Nusskern, D.B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, A. Yao, J. Ye, M. Zhan, W. Zhang, H. Zhang, Q. Zhao, L. Zheng, F. Zhong, W. Zhong, S. Zhu, S. Zhao, D. Gilbert, S. Baumhueter, G. Spier, C. Carter, A. Cravchik, T. Woodage, F. Ali, H. An, A. Awe, D. Baldwin, H. Baden, M. Barnstead, I. Barrow, K. Beeson, D. Busam, A. Carver, A. Center, M.L. Cheng, L. Curry, S. Danaher, L. Davenport, R. Desilets, S. Dietz, K. Dodson, L. Doup, S. Ferriera, N. Garg, A. Gluecksmann, B. Hart, J. Haynes, C. Haynes, C. Heiner, S. Hladun, D. Hostin, J. Houck, T. Howland, C. Ibegwam, J. Johnson, F. Kalush, L. Kline, S. Koduru, A. Love, F. Mann, D. May, S. McCawley, T. McIntosh, I. McMullen, M. Moy, L. Moy, B. Murphy, K. Nelson, C. Pfannkoch, E. Pratts, V. Puri, H. Qureshi, M. Reardon, R. Rodriguez, Y.H. Rogers, D. Romblad, B. Ruhfel, R. Scott, C. Sitter, M. Smallwood, E. Stewart, R. Strong, E. Suh, R. Thomas, N.N. Tint, S. Tse, C. Vech, G. Wang, J. Wetter, S. Williams, M. Williams, S. Windsor, E. Winn-Deen, K. Wolfe, J. Zaveri, K. Zaveri, J.F. Abril, R. Guigo, M.J. Campbell, K.V. Sjolander, B. Karlak, A. Kejariwal, H. Mi, B. Lazareva, T. Hatton, A. Narechania, K. Diemer, A. Muruganujan, N. Guo, S. Sato, V. Bafna, S. Istrail, R. Lippert, R. Schwartz, B. Walenz, S. Yooseph, D. Allen, A. Basu, J. Baxendale, L. Blick, M. Caminha, J. Carnes-Stine, P. Caulk, Y.H. Chiang, M. Coyne, C. Dahlke, A. Mays, M. Dombroski, M. Donnelly, D. Ely, S. Esparham, C. Fosler, H. Gire, S. Glanowski, K. Glasser, A. Glodek, M. Gorokhov, K. Graham, B. Gropman, M. Harris, J. Heil, S. Henderson, J. Hoover, D. Jennings, C. Jordan, J. Jordan, J. Kasha, L. Kagan, C. Kraft, A. Levitsky, M. Lewis, X. Liu, J. Lopez, D. Ma, W. Majoros, J. McDaniel, S. Murphy, M. Newman, T. Nguyen, N. Nguyen, M. Nodell, S. Pan, J. Peck, M. Peterson, W. Rowe, R. Sanders, J. Scott, M. Simpson, T. Smith, A. Sprague, T. Stockwell, R. Turner, E. Venter, M. Wang, M. Wen, D. Wu, M. Wu, A. Xia, A. Zandieh and X. Zhu, *The sequence of the human genome.* Science, 2001. **291**(5507): p. 1304-51.

14. UniProt_Consortium, *UniProt: a hub for protein information.* Nucleic Acids Research, 2015. **43**(D1): p. D204-D212.

15. Kim, M.S., S.M. Pinto, D. Getnet, R.S. Nirujogi, S.S. Manda, R. Chaerkady, A.K. Madugundu, D.S. Kelkar, R. Isserlin, S. Jain, J.K. Thomas, B. Muthusamy, P. Leal-Rojas, P. Kumar, N.A. Sahasrabuddhe, L. Balakrishnan, J. Advani, B. George, S. Renuse, L.D. Selvan, A.H. Patil, V. Nanjappa, A. Radhakrishnan, S. Prasad, T. Subbannayya, R. Raju, M. Kumar, S.K. Sreenivasamurthy, A. Marimuthu, G.J. Sathe, S. Chavan, K.K. Datta, Y. Subbannayya, A. Sahu, S.D. Yelamanchi, S. Jayaram, P. Rajagopalan, J. Sharma, K.R. Murthy, N. Syed, R. Goel, A.A. Khan, S. Ahmad, G. Dey, K. Mudgal, A. Chatterjee, T.C. Huang, J. Zhong, X. Wu,

P.G. Shaw, D. Freed, M.S. Zahari, K.K. Mukherjee, S. Shankar, A. Mahadevan, H. Lam, C.J. Mitchell, S.K. Shankar, P. Satishchandra, J.T. Schroeder, R. Sirdeshmukh, A. Maitra, S.D. Leach, C.G. Drake, M.K. Halushka, T.S. Prasad, R.H. Hruban, C.L. Kerr, G.D. Bader, C.A. Iacobuzio-Donahue, H. Gowda, and A. Pandey, *A draft map of the human proteome.* Nature, 2014. **509**(7502): p. 575-81.

16. Wilhelm, M., J. Schlegl, H. Hahne, A. Moghaddas Gholami, M. Lieberenz, M.M. Savitski, E. Ziegler, L. Butzmann, S. Gessulat, H. Marx, T. Mathieson, S. Lemeer, K. Schnatbaum, U. Reimer, H. Wenschuh, M. Mollenhauer, J. Slotta-Huspenina, J.H. Boese, M. Bantscheff, A. Gerstmair, F. Faerber, and B. Kuster, *Mass-spectrometry-based draft of the human proteome.* Nature, 2014. **509**(7502): p. 582-7.
17. Uhlen, M., L. Fagerberg, B.M. Hallstrom, C. Lindskog, P. Oksvold, A. Mardinoglu, A. Sivertsson, C. Kampf, E. Sjostedt, A. Asplund, I. Olsson, K. Edlund, E. Lundberg, S. Navani, C.A. Szigyarto, J. Odeberg, D. Djureinovic, J.O. Takanen, S. Hober, T. Alm, P.H. Edqvist, H. Berling, H. Tegel, J. Mulder, J. Rockberg, P. Nilsson, J.M. Schwenk, M. Hamsten, K. von Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. von Heijne, J. Nielsen, and F. Ponten, *Proteomics. Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.
18. Perez-Riverol, Y., E. Alpi, R. Wang, H. Hermjakob, and J.A. Vizcaino, *Making proteomics data accessible and reusable: current state of proteomics databases and repositories.* Proteomics, 2015. **15**(5-6): p. 930-49.
19. Smith, L.M., N.L. Kelleher, and P. Consortium for Top Down, *Proteoform: a single term describing protein complexity.* Nat Methods, 2013. **10**(3): p. 186-7.
20. Johnson, H. and C.E. Eyers, *Analysis of post-translational modifications by LC-MS/MS.* Methods Mol Biol, 2010. **658**: p. 93-108.
21. Silva, A.M., R. Vitorino, M.R. Domingues, C.M. Spickett, and P. Domingues, *Post-translational modifications and mass spectrometry detection.* Free Radic Biol Med, 2013. **65**: p. 925-41.
22. Yang, C., X. Zhong, and L. Li, *Recent advances in enrichment and separation strategies for mass spectrometry-based phosphoproteomics.* Electrophoresis, 2014. **35**(24): p. 3418-29.
23. Zhao, Y. and O.N. Jensen, *Modification-specific proteomics: Strategies for characterization of post-translational modifications using enrichment techniques.* Proteomics, 2009. **9**(20): p. 4632-4641.
24. Berglund, L., E. Bjorling, K. Jonasson, J. Rockberg, L. Fagerberg, C. Al-Khalili Szigyarto, A. Sivertsson, and M. Uhlen, *A whole-genome bioinformatics approach to selection of antigens for systematic antibody generation.* Proteomics, 2008. **8**(14): p. 2832-9.
25. Bordeaux, J., A. Welsh, S. Agarwal, E. Killiam, M. Baquero, J. Hanna, V. Anagnostou, and D. Rimm, *Antibody validation.* Biotechniques, 2010. **48**(3): p. 197-209.
26. Bradbury, A. and A. Pluckthun, *Standardize antibodies used in research.* Nature, 2015. **518**(7537): p. 27-9.
27. Couchman, J.R., *Commercial antibodies: the good, bad, and really ugly.* J Histochem Cytochem, 2009. **57**(1): p. 7-8.

28. Coons, A.H., H.J. Creech, and R.N. Jones, *Immunological Properties of an Antibody Containing a Fluorescent Group.* Experimental Biology and Medicine, 1941. **47**(2): p. 200-202.
29. Yalow, R.S. and S.A. Berson, *Assay of plasma insulin in human subjects by immunological methods.* Nature, 1959. **184 (Suppl 21)**: p. 1648-9.
30. Engvall, E. and P. Perlmann, *Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G.* Immunochemistry, 1971. **8**(9): p. 871-4.
31. Kohler, G. and C. Milstein, *Continuous cultures of fused cells secreting antibody of predefined specificity. 1975.* J Immunol, 1975. **174**(5): p. 2453-5.
32. Burnette, W.N., *"Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A.* Anal Biochem, 1981. **112**(2): p. 195-203.
33. Phelan, M.L. and S. Nock, *Generation of bioreagents for protein chips.* Proteomics, 2003. **3**(11): p. 2123-34.
34. Uhlen, M. and F. Ponten, *Antibody-based proteomics for human tissue profiling.* Mol Cell Proteomics, 2005. **4**(4): p. 384-93.
35. Geisse, S., H. Gram, B. Kleuser, and H.P. Kocher, *Eukaryotic expression systems: a comparison.* Protein Expr Purif, 1996. **8**(3): p. 271-82.
36. Hancock, D.C. and N.J. O'Reilly, *Synthetic peptides as antigens for antibody production.* Methods Mol Biol, 2005. **295**: p. 13-26.
37. Hjelm, B., C.D. Fernandez, J. Lofblom, S. Stahl, H. Johannesson, J. Rockberg, and M. Uhlen, *Exploring epitopes of antibodies toward the human tryptophanyl-tRNA synthetase.* N Biotechnol, 2010. **27**(2): p. 129-37.
38. Colwill, K., H. Persson, N.E. Jarvik, A. Wyrzucki, J. Wojcik, A. Koide, A.A. Kossiakoff, S. Koide, S. Sidhu, M.R. Dyson, K. Pershad, J.D. Pavlovic, A. Karatt-Vellatt, D.J. Schofield, B.K. Kay, J. McCafferty, M. Mersmann, D. Meier, J. Mersmann, S. Helmsing, M. Hust, S. Dubel, S. Berkowicz, A. Freemantle, M. Spiegel, A. Sawyer, D. Layton, E. Nice, A. Dai, O. Rocks, K. Williton, F.A. Fellouse, K. Hersi, T. Pawson, P. Nilsson, M. Sundberg, R. Sjoberg, A. Sivertsson, J.M. Schwenk, J.O. Takanen, S. Hober, M. Uhlen, L.G. Dahlgren, A. Flores, I. Johansson, J. Weigelt, L. Crombet, P. Loppnau, I. Kozieradzki, D. Cossar, C.H. Arrowsmith, A.M. Edwards, and S. Graslund, *A roadmap to generate renewable protein binders to the human proteome.* Nat Methods, 2011.
39. De Masi, F., P. Chiarella, H. Wilhelm, M. Massimi, B. Bullard, W. Ansorge, and A. Sawyer, *High throughput production of mouse monoclonal antibodies using antigen microarrays.* Proteomics, 2005. **5**(16): p. 4070-81.
40. Mersmann, M., D. Meier, J. Mersmann, S. Helmsing, P. Nilsson, S. Graslund, K. Colwill, M. Hust, and S. Dubel, *Towards proteome scale antibody selections using phage display.* N Biotechnol, 2010. **27**(2): p. 118-28.
41. Nilsson, P., L. Paavilainen, K. Larsson, J. Odling, M. Sundberg, A.C. Andersson, C. Kampf, A. Persson, C. Al-Khalili Szigyarto, J. Ottosson, E. Bjorling, S. Hober, H. Wernerus, K. Wester, F. Ponten, and M. Uhlen,

Towards a human proteome atlas: high-throughput generation of mono-specific antibodies for tissue profiling. Proteomics, 2005. **5**(17): p. 4327-37.

42. Pershad, K., J.D. Pavlovic, S. Graslund, P. Nilsson, K. Colwill, A. Karatt-Vellatt, D.J. Schofield, M.R. Dyson, T. Pawson, B.K. Kay, and J. McCafferty, *Generating a panel of highly specific antibodies to 20 human SH2 domains by phage display.* Protein Eng Des Sel, 2010. **23**(4): p. 279-88.
43. Abbas, A.K. and A.H. Lichtman, *Basic immunology : functions and disorders of the immune system*. 2011, Philadelphia, Pa: Saunders.
44. Woof, J.M. and D.R. Burton, *Human antibody-Fc receptor interactions illuminated by crystal structures.* Nat Rev Immunol, 2004. **4**(2): p. 89-99.
45. Wingren, C., C. Steinhauer, J. Ingvarsson, E. Persson, K. Larsson, and C.A. Borrebaeck, *Microarrays based on affinity-tagged single-chain Fv antibodies: sensitive detection of analyte in complex proteomes.* Proteomics, 2005. **5**(5): p. 1281-91.
46. Lipman, N.S., L.R. Jackson, L.J. Trudel, and F. Weis-Garcia, *Monoclonal versus polyclonal antibodies: distinguishing characteristics, applications, and information resources.* ILAR J, 2005. **46**(3): p. 258-68.
47. Rockberg, J., J.M. Schwenk, and M. Uhlen, *Discovery of epitopes for targeting the human epidermal growth factor receptor 2 (HER2) with antibodies.* Mol Oncol, 2009.
48. Angenendt, P., J. Glokler, Z. Konthur, H. Lehrach, and D.J. Cahill, *3D protein microarrays: performing multiplex immunoassays on a single chip.* Anal Chem, 2003. **75**(17): p. 4368-72.
49. Wide, L., H. Bennich, and S.G. Johansson, *Diagnosis of allergy by an in-vitro test for allergen antibodies.* Lancet, 1967. **2**(7526): p. 1105-7.
50. Warford, A., W. Howat, and J. McCafferty, *Expression profiling by high-throughput immunohistochemistry.* J Immunol Methods, 2004. **290**(1-2): p. 81-92.
51. Porter, R.R., *The hydrolysis of rabbit y-globulin and antibodies with crystalline papain.* Biochem J, 1959. **73**: p. 119-26.
52. Bird, R.E., K.D. Hardman, J.W. Jacobson, S. Johnson, B.M. Kaufman, S.M. Lee, T. Lee, S.H. Pope, G.S. Riordan, and M. Whitlow, *Single-chain antigen-binding proteins.* Science, 1988. **242**(4877): p. 423-6.
53. McCafferty, J., A.D. Griffiths, G. Winter, and D.J. Chiswell, *Phage antibodies: filamentous phage displaying antibody variable domains.* Nature, 1990. **348**(6301): p. 552-4.
54. Pansri, P., N. Jaruseranee, K. Rangnoi, P. Kristensen, and M. Yamabhai, *A compact phage display human scFv library for selection of antibodies to a wide variety of antigens.* BMC Biotechnol, 2009. **9**: p. 6.
55. Kwong, K.Y., S. Baskar, H. Zhang, C.L. Mackall, and C. Rader, *Generation, affinity maturation, and characterization of a human anti-human NKG2D monoclonal antibody with dual antagonistic and agonistic activity.* J Mol Biol, 2008. **384**(5): p. 1143-56.
56. Ayoglu, B., A. Haggmark, M. Neiman, U. Igel, M. Uhlen, J.M. Schwenk, and P. Nilsson, *Systematic antibody and antigen-based proteomic profiling with microarrays.* Expert Rev Mol Diagn, 2011. **11**(2): p. 219-34.

57. Lin, H. and V.W. Cornish, *Screening and selection methods for large-scale analysis of protein function.* Angew Chem Int Ed Engl, 2002. **41**(23): p. 4402-25.
58. Better, M., C.P. Chang, R.R. Robinson, and A.H. Horwitz, *Escherichia coli secretion of an active chimeric antibody fragment.* Science, 1988. **240**(4855): p. 1041-3.
59. Skerra, A. and A. Pluckthun, *Assembly of a functional immunoglobulin Fv fragment in Escherichia coli.* Science, 1988. **240**(4855): p. 1038-41.
60. Fernandez, L.A., *Prokaryotic expression of antibodies and affibodies.* Curr Opin Biotechnol, 2004. **15**(4): p. 364-73.
61. Nord, K., E. Gunneriusson, J. Ringdahl, S. Stahl, M. Uhlen, and P.A. Nygren, *Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain.* Nat Biotechnol, 1997. **15**(8): p. 772-7.
62. Muyldermans, S., *Nanobodies: natural single-domain antibodies.* Annu Rev Biochem, 2013. **82**: p. 775-97.
63. Muyldermans, S., T. Atarhouch, J. Saldanha, J.A. Barbosa, and R. Hamers, *Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains.* Protein Eng, 1994. **7**(9): p. 1129-35.
64. Helma, J., M.C. Cardoso, S. Muyldermans, and H. Leonhardt, *Nanobodies and recombinant binders in cell biology.* J Cell Biol, 2015. **209**(5): p. 633-644.
65. Uhlen, M., P. Oksvold, L. Fagerberg, E. Lundberg, K. Jonasson, M. Forsberg, M. Zwahlen, C. Kampf, K. Wester, S. Hober, H. Wernerus, L. Bjorling, and F. Ponten, *Towards a knowledge-based Human Protein Atlas.* Nat Biotechnol, 2010. **28**(12): p. 1248-50.
66. Berglund, L., E. Bjorling, P. Oksvold, L. Fagerberg, A. Asplund, C.A. Szigyarto, A. Persson, J. Ottosson, H. Wernerus, P. Nilsson, E. Lundberg, A. Sivertsson, S. Navani, K. Wester, C. Kampf, S. Hober, F. Ponten, and M. Uhlen, *A genecentric Human Protein Atlas for expression profiles based on antibodies.* Mol Cell Proteomics, 2008. **7**(10): p. 2019-27.
67. Uhlen, M., E. Bjorling, C. Agaton, C.A. Szigyarto, B. Amini, E. Andersen, A.C. Andersson, P. Angelidou, A. Asplund, C. Asplund, L. Berglund, K. Bergstrom, H. Brumer, D. Cerjan, M. Ekstrom, A. Elobeid, C. Eriksson, L. Fagerberg, R. Falk, J. Fall, M. Forsberg, M.G. Bjorklund, K. Gumbel, A. Halimi, I. Hallin, C. Hamsten, M. Hansson, M. Hedhammar, G. Hercules, C. Kampf, K. Larsson, M. Lindskog, W. Lodewyckx, J. Lund, J. Lundeberg, K. Magnusson, E. Malm, P. Nilsson, J. Odling, P. Oksvold, I. Olsson, E. Oster, J. Ottosson, L. Paavilainen, A. Persson, R. Rimini, J. Rockberg, M. Runeson, A. Sivertsson, A. Skollermo, J. Steen, M. Stenvall, F. Sterky, S. Stromberg, M. Sundberg, H. Tegel, S. Tourle, E. Wahlund, A. Walden, J. Wan, H. Wernerus, J. Westberg, K. Wester, U. Wrethagen, L.L. Xu, S. Hober, and F. Ponten, *A human protein atlas for normal and cancer tissues based on antibody proteomics.* Mol Cell Proteomics, 2005. **4**(12): p. 1920-32.
68. Porath, J., J. Carlsson, I. Olsson, and G. Belfrage, *Metal chelate affinity chromatography, a new approach to protein fractionation.* Nature, 1975. **258**(5536): p. 598-9.

69. Sjolander, A., P.A. Nygren, S. Stahl, K. Berzins, M. Uhlen, P. Perlmann, and R. Andersson, *The serum albumin-binding region of streptococcal protein G: a bacterial fusion partner with carrier-related properties.* J Immunol Methods, 1997. **201**(1): p. 115-23.
70. Agaton, C., J. Galli, I. Hoiden Guthenberg, L. Janzon, M. Hansson, A. Asplund, E. Brundell, S. Lindberg, I. Ruthberg, K. Wester, D. Wurtz, C. Hoog, J. Lundeberg, S. Stahl, F. Ponten, and M. Uhlen, *Affinity proteomics for systematic protein profiling of chromosome 21 gene products in human tissues.* Mol Cell Proteomics, 2003. **2**(6): p. 405-14.
71. Tegel, H., J. Steen, A. Konrad, H. Nikdin, K. Pettersson, M. Stenvall, S. Tourle, U. Wrethagen, L. Xu, L. Yderland, M. Uhlen, S. Hober, and J. Ottosson, *High-throughput protein production--lessons from scaling up from 10 to 288 recombinant proteins per week.* Biotechnol J, 2009. **4**(1): p. 51-7.
72. Steen, J., M. Uhlen, S. Hober, and J. Ottosson, *High-throughput protein purification using an automated set-up for high-yield affinity chromatography.* Protein Expr Purif, 2006. **46**(2): p. 173-8.
73. Lundberg, E. and M. Uhlen, *Creation of an antibody-based subcellular protein atlas.* Proteomics, 2010. **10**(22): p. 3984-96.
74. Baker, M., *Blame it on the antibodies.* Nature, 2015. **521**(7552): p. 274-6.
75. Begley, C.G. and L.M. Ellis, *Raise standards for preclinical cancer research.* Nature, 2012. **483**(7391): p. 531-3.
76. Michel, M.C., T. Wieland, and G. Tsujimoto, *How reliable are G-protein-coupled receptor antibodies?* Naunyn Schmiedebergs Arch Pharmacol, 2009. **379**(4): p. 385-8.
77. Egelhofer, T.A., A. Minoda, S. Klugman, K. Lee, P. Kolasinska-Zwierz, A.A. Alekseyenko, M.S. Cheung, D.S. Day, S. Gadel, A.A. Gorchakov, T. Gu, P.V. Kharchenko, S. Kuan, I. Latorre, D. Linder-Basso, Y. Luu, Q. Ngo, M. Perry, A. Rechtsteiner, N.C. Riddle, Y.B. Schwartz, G.A. Shanower, A. Vielle, J. Ahringer, S.C. Elgin, M.I. Kuroda, V. Pirrotta, B. Ren, S. Strome, P.J. Park, G.H. Karpen, R.D. Hawkins, and J.D. Lieb, *An assessment of histone-modification antibody quality.* Nat Struct Mol Biol, 2011. **18**(1): p. 91-3.
78. Voskuil, J.L.A., *Commercial antibodies and their validation.* F1000Research, 2014. **3**: p. 232.
79. Bradbury, A.M. and A. Pluckthun, *Validate recombinants once.* Nature, 2015. **520**(7547): p. 295-295.
80. Freedman, L.P., *Validate recombinants too.* Nature, 2015. **518**(7540): p. 483-483.
81. Polakiewicz, R.D., *The solution is validation.* Nature, 2015. **518**(7540): p. 483-483.
82. James, L.C. and D.S. Tawfik, *The specificity of cross-reactivity: promiscuous antibody binding involves specific hydrogen bonds rather than nonspecific hydrophobic stickiness.* Protein Sci, 2003. **12**(10): p. 2183-93.
83. Olsson, N., S. Wallin, P. James, C.A.K. Borrebaeck, and C. Wingren, *Epitope-specificity of recombinant antibodies reveals promiscuous peptide-binding properties.* Protein Science, 2012. **21**(12): p. 1897-1910.

84. Bostrom, T., H.J. Johansson, J. Lehtio, M. Uhlen, and S. Hober, *Investigating the applicability of antibodies generated within the human protein atlas as capture agents in immunoenrichment coupled to mass spectrometry.* J Proteome Res, 2014. **13**(10): p. 4424-35.
85. Edfors, F., T. Bostrom, B. Forsstrom, M. Zeiler, H. Johansson, E. Lundberg, S. Hober, J. Lehtio, M. Mann, and M. Uhlen, *Immunoproteomics using polyclonal antibodies and stable isotope-labeled affinity-purified recombinant proteins.* Mol Cell Proteomics, 2014. **13**(6): p. 1611-24.
86. Tansey, W.P., *Nondenaturing Protein Immunoprecipitation from Mammalian Cells.* Cold Spring Harbor Protocols, 2007. **2007**(1): p. pdb.prot4620.
87. Tansey, W.P., *Denaturing Protein Immunoprecipitation from Mammalian Cells.* Cold Spring Harbor Protocols, 2007. **2007**(1): p. pdb.prot4619.
88. Schwenk, J.M., U. Igel, M. Neiman, H. Langen, C. Becker, A. Bjartell, F. Ponten, F. Wiklund, H. Gronberg, P. Nilsson, and M. Uhlen, *Toward next generation plasma profiling via heat-induced epitope retrieval and array-based assays.* Mol Cell Proteomics, 2010. **9**(11): p. 2497-507.
89. Marcon, E., H. Jain, A. Bhattacharya, H. Guo, S. Phanse, S. Pu, G. Byram, B.C. Collins, E. Dowdell, M. Fenner, X. Guo, A. Hutchinson, J.J. Kennedy, B. Krastins, B. Larsen, Z.Y. Lin, M.F. Lopez, P. Loppnau, S. Miersch, T. Nguyen, J.B. Olsen, M. Paduch, M. Ravichandran, A. Seitova, G. Vadali, M.S. Vogelsang, J.R. Whiteaker, G. Zhong, N. Zhong, L. Zhao, R. Aebersold, C.H. Arrowsmith, A. Emili, L. Frappier, A.C. Gingras, M. Gstaiger, A.G. Paulovich, S. Koide, A.A. Kossiakoff, S.S. Sidhu, S.J. Wodak, S. Graslund, J.F. Greenblatt, and A.M. Edwards, *Assessment of a method to characterize antibody selectivity and specificity for use in immunoprecipitation.* Nat Methods, 2015.
90. Fodor, S.P., J.L. Read, M.C. Pirrung, L. Stryer, A.T. Lu, and D. Solas, *Light-directed, spatially addressable parallel chemical synthesis.* Science, 1991. **251**(4995): p. 767-73.
91. Schena, M., D. Shalon, R.W. Davis, and P.O. Brown, *Quantitative monitoring of gene expression patterns with a complementary DNA microarray.* Science, 1995. **270**(5235): p. 467-70.
92. Ekins, R.P., *Multi-analyte immunoassay.* J Pharm Biomed Anal, 1989. **7**(2): p. 155-68.
93. Ekins, R.P., *Ligand assays: from electrophoresis to miniaturized microarrays.* Clin Chem, 1998. **44**(9): p. 2015-30.
94. MacBeath, G. and S.L. Schreiber, *Printing proteins as microarrays for high-throughput function determination.* Science, 2000. **289**(5485): p. 1760-3.
95. Templin, M.F., D. Stoll, J.M. Schwenk, O. Potz, S. Kramer, and T.O. Joos, *Protein microarrays: promising tools for proteomic research.* Proteomics, 2003. **3**(11): p. 2155-66.
96. Seong, S.Y. and C.Y. Choi, *Current status of protein chip development in terms of fabrication and application.* Proteomics, 2003. **3**(11): p. 2176-89.
97. Chandra, H., P.J. Reddy, and S. Srivastava, *Protein microarrays and novel detection platforms.* Expert Rev Proteomics, 2011. **8**(1): p. 61-79.
98. Melioli, G., F. Bonifazi, S. Bonini, E. Maggi, M. Mussap, G. Passalacqua, E.R. Rossi, A. Vacca, and G.W. Canonica, *The ImmunoCAP ISAC*

molecular allergology approach in adult multi-sensitized Italian patients with respiratory symptoms. Clin Biochem, 2011.

99. Moreira, P.F., K. Gangl, A. Vieira Fde, L.H. Ynoue, B. Linhart, S. Flicker, H. Fiebig, I. Swoboda, M. Focke-Tejkl, E.A. Taketomi, R. Valenta, and V. Niederberger, *Allergen Microarray Indicates Pooideae Sensitization in Brazilian Grass Pollen Allergic Patients.* PLoS One, 2015. **10**(6): p. e0128402.
100. Hansson, M., L. Mathsson, T. Schlederer, L. Israelsson, P. Matsson, L. Nogueira, P.J. Jakobsson, K. Lundberg, V. Malmstrom, G. Serre, R. Holmdahl, M. Nystrand, L. Klareskog, and J. Ronnelid, *Validation of a multiplex chip-based assay for the detection of autoantibodies against citrullinated peptides.* Arthritis Res Ther, 2012. **14**(5): p. R201.
101. Hueber, W., B.H. Tomooka, F. Batliwalla, W. Li, P.A. Monach, R.J. Tibshirani, R.F. Van Vollenhoven, J. Lampa, K. Saito, Y. Tanaka, M.C. Genovese, L. Klareskog, P.K. Gregersen, and W.H. Robinson, *Blood autoantibody and cytokine profiles predict response to anti-tumor necrosis factor therapy in rheumatoid arthritis.* Arthritis Res Ther, 2009. **11**(3): p. R76.
102. Rosenberg, J.M. and P.J. Utz, *Protein microarrays: a new tool for the study of autoantibodies in immunodeficiency.* Front Immunol, 2015. **6**: p. 138.
103. Abel, L., S. Kutschki, M. Turewicz, M. Eisenacher, J. Stoutjesdijk, H.E. Meyer, D. Woitalla, and C. May, *Autoimmune profiling with protein microarrays in clinical applications.* Biochim Biophys Acta, 2014. **1844**(5): p. 977-87.
104. Crompton, P.D., M.A. Kayala, B. Traore, K. Kayentao, A. Ongoiba, G.E. Weiss, D.M. Molina, C.R. Burk, M. Waisberg, A. Jasinskas, X. Tan, S. Doumbo, D. Doumtabe, Y. Kone, D.L. Narum, X. Liang, O.K. Doumbo, L.H. Miller, D.L. Doolan, P. Baldi, P.L. Felgner, and S.K. Pierce, *A prospective analysis of the Ab response to Plasmodium falciparum before and after a malaria season by protein microarray.* Proc Natl Acad Sci U S A, 2010. **107**(15): p. 6958-63.
105. Xu, Y., J.F. Bruno, and B.J. Luft, *Profiling the humoral immune response to Borrelia burgdorferi infection with protein microarrays.* Microb Pathog, 2008. **45**(5-6): p. 403-7.
106. Viswanathan, K., N. Kilcullen, C. Morrell, S.J. Thistlethwaite, M.U. Sivananthan, T.B. Hassan, J.H. Barth, and A.S. Hall, *Heart-type fatty acid-binding protein predicts long-term mortality and re-infarction in consecutive patients with suspected acute coronary syndrome who are troponin-negative.* J Am Coll Cardiol, 2010. **55**(23): p. 2590-8.
107. Borrebaeck, C.A. and C. Wingren, *Design of high-density antibody microarrays for disease proteomics: key technological issues.* J Proteomics, 2009. **72**(6): p. 928-35.
108. Cretich, M., F. Damin, and M. Chiari, *Protein microarray technology: how far off is routine diagnostics?* Analyst, 2014. **139**(3): p. 528-42.
109. Schroder, C., A. Jacob, S. Tonack, T.P. Radon, M. Sill, M. Zucknick, S. Ruffer, E. Costello, J.P. Neoptolemos, T. Crnogorac-Jurcevic, A. Bauer, K. Fellenberg, and J.D. Hoheisel, *Dual-color proteomic profiling of complex samples with a microarray of 810 cancer-related antibodies.* Mol Cell Proteomics, 2010. **9**(6): p. 1271-80.

110. Chiari, M., M. Cretich, A. Corti, F. Damin, G. Pirri, and R. Longhi, *Peptide microarrays for the characterization of antigenic regions of human chromogranin A.* Proteomics, 2005. **5**(14): p. 3600-3.
111. Forsstrom, B., B.B. Axnas, K.P. Stengele, J. Buhler, T.J. Albert, T.A. Richmond, F.J. Hu, P. Nilsson, E.P. Hudson, J. Rockberg, and M. Uhlen, *Proteome-wide epitope mapping of antibodies using ultra-dense peptide arrays.* Mol Cell Proteomics, 2014. **13**(6): p. 1585-97.
112. Janzi, M., J. Odling, Q. Pan-Hammarstrom, M. Sundberg, J. Lundeberg, M. Uhlen, L. Hammarstrom, and P. Nilsson, *Serum microarrays for large scale screening of protein levels.* Mol Cell Proteomics, 2005. **4**(12): p. 1942-7.
113. Pawlak, M., E. Schick, M.A. Bopp, M.J. Schneider, P. Oroszlan, and M. Ehrat, *Zeptosens' protein microarrays: a novel high performance microarray platform for low abundance protein analysis.* Proteomics, 2002. **2**(4): p. 383-93.
114. Romanov, V., S.N. Davidoff, A.R. Miles, D.W. Grainger, B.K. Gale, and B.D. Brooks, *A critical comparison of protein microarray fabrication technologies.* Analyst, 2014. **139**(6): p. 1303-26.
115. Angenendt, P., *Progress in protein and antibody microarray technology.* Drug Discov Today, 2005. **10**(7): p. 503-11.
116. Kusnezow, W. and J.D. Hoheisel, *Solid supports for microarray immunoassays.* J Mol Recognit, 2003. **16**(4): p. 165-76.
117. Sanchez-Carbayo, M., *Antibody arrays: technical considerations and clinical applications in cancer.* Clin Chem, 2006. **52**(9): p. 1651-9.
118. Usui-Aoki, K., K. Shimada, and H. Koga, *A novel antibody microarray format using non-covalent antibody immobilization with chemiluminescent detection.* Mol Biosyst, 2007. **3**(1): p. 36-42.
119. Haab, B.B., *Methods and applications of antibody microarrays in cancer research.* Proteomics, 2003. **3**(11): p. 2116-22.
120. Wingren, C. and C.A. Borrebaeck, *Antibody microarray analysis of directly labelled complex proteomes.* Curr Opin Biotechnol, 2008. **19**(1): p. 55-61.
121. Kusnezow, W., V. Banzon, C. Schroder, R. Schaal, J.D. Hoheisel, S. Ruffer, P. Luft, A. Duschl, and Y.V. Syagailo, *Antibody microarray-based profiling of complex specimens: systematic evaluation of labeling strategies.* Proteomics, 2007. **7**(11): p. 1786-99.
122. Wingren, C., J. Ingvarsson, L. Dexlin, D. Szul, and C.A. Borrebaeck, *Design of recombinant antibody microarrays for complex proteome analysis: choice of sample labeling-tag and solid support.* Proteomics, 2007. **7**(17): p. 3055-65.
123. Schwenk, J.M., J. Lindberg, M. Sundberg, M. Uhlen, and P. Nilsson, *Determination of binding specificities in highly multiplexed bead-based assays for antibody proteomics.* Mol Cell Proteomics, 2007. **6**(1): p. 125-32.
124. Andersson, P., G. Jesson, G. Kylberg, G. Ekstrand, and G. Thorsen, *Parallel nanoliter microfluidic analysis system.* Anal Chem, 2007. **79**(11): p. 4022-30.
125. Eriksson, C., C. Agaton, R. Kange, M. Sundberg, P. Nilsson, B. Ek, M. Uhlen, M. Gustafsson, and S. Hober, *Microfluidic analysis of antibody*

specificity in a compact disk format. J Proteome Res, 2006. **5**(7): p. 1568-74.
126. de Hoffmann, E. and V. Stroobant, *Mass Spectrometry*. 2007, Chichester, England: John Wiley &Sons.
127. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons.* Anal Chem, 1988. **60**(20): p. 2299-301.
128. Tanaka, K., H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, and T. Matsuo, *Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry.* Rapid Commun. Mass Spectrom., 1988. **2**(8): p. 151-1533.
129. Yamashita, M. and J.B. Fenn, *Electrospray ion source. Another variation on the free-jet theme.* The Journal of Physical Chemistry, 1984. **88**(20): p. 4451-4459.
130. Wilm, M., *Principles of electrospray ionization.* Mol Cell Proteomics, 2011. **10**(7): p. M111 009407.
131. Miller, P.E. and M.B. Denton, *The quadrupole mass filter: Basic operating concepts.* Journal of Chemical Education, 1986. **63**(7): p. 617-622.
132. March, R.E., *An Introduction to Quadrupole Ion Trap Mass Spectrometry.* Journal of Mass Spectrometry, 1997. **32**(4): p. 351-369.
133. Mamyrin, B.A., *Time-of-flight mass spectrometry (concepts, achievements, and prospects).* International Journal of Mass Spectrometry, 2001. **206**(3): p. 251-266.
134. Comisarow, M.B. and A.G. Marshall, *Fourier transform ion cyclotron resonance spectroscopy.* Chemical Physics Letters, 1974. **25**(2): p. 282-283.
135. Makarov, A., *Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis.* Anal Chem, 2000. **72**(6): p. 1156-62.
136. Glish, G.L. and D.J. Burinsky, *Hybrid mass spectrometers for tandem mass spectrometry.* J Am Soc Mass Spectrom, 2008. **19**(2): p. 161-72.
137. Scigelova, M. and A. Makarov, *Orbitrap mass analyzer--overview and applications in proteomics.* Proteomics, 2006. **6 Suppl 2**: p. 16-21.
138. Michalski, A., E. Damoc, J.P. Hauschild, O. Lange, A. Wieghaus, A. Makarov, N. Nagaraj, J. Cox, M. Mann, and S. Horning, *Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer.* Mol Cell Proteomics, 2011. **10**(9): p. M111 011015.
139. Horvath, C.G., B.A. Preiss, and S.R. Lipsky, *Fast liquid chromatography: an investigation of operating parameters and the separation of nucleotides on pellicular ion exchangers.* Anal Chem, 1967. **39**(12): p. 1422-8.
140. Yates, J.R., C.I. Ruse, and A. Nakorchevsky, *Proteomics by mass spectrometry: approaches, advances, and applications.* Annu Rev Biomed Eng, 2009. **11**: p. 49-79.
141. Mischak, H., J.J. Coon, J. Novak, E.M. Weissinger, J.P. Schanstra, and A.F. Dominiczak, *Capillary electrophoresis-mass spectrometry as a powerful tool in biomarker discovery and clinical diagnosis: an update of recent developments.* Mass Spectrom Rev, 2009. **28**(5): p. 703-24.

142. Fenn, J.B., M. Mann, C.K. Meng, S.F. Wong, and C.M. Whitehouse, *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64-71.
143. Iribarne, J.V. and B.A. Thomson, *On the evaporation of small ions from charged droplets.* The Journal of Chemical Physics, 1976. **64**(6): p. 2287-2294.
144. Dole, M., L.L. Mack, R.L. Hines, R.C. Mobley, L.D. Ferguson, and M.B. Alice, *Molecular Beams of Macroions.* The Journal of Chemical Physics, 1968. **49**(5): p. 2240-2249.
145. Geiger, T., J. Cox, and M. Mann, *Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation.* Mol Cell Proteomics, 2010. **9**(10): p. 2252-61.
146. Catherman, A.D., O.S. Skinner, and N.L. Kelleher, *Top Down proteomics: facts and perspectives.* Biochem Biophys Res Commun, 2014. **445**(4): p. 683-93.
147. Forbes, A.J., M.T. Mazur, H.M. Patel, C.T. Walsh, and N.L. Kelleher, *Toward efficient analysis of >70 kDa proteins with 100% sequence coverage.* Proteomics, 2001. **1**(8): p. 927-33.
148. Tsiatsiani, L. and A.J. Heck, *Proteomics beyond trypsin.* FEBS J, 2015.
149. Lu, B. and T. Chen, *Algorithms for de novo peptide sequencing using tandem mass spectrometry.* Drug Discovery Today: BIOSILICO, 2004. **2**(2): p. 85-90.
150. Eng, J.K., A.L. McCormack, and J.R. Yates, *An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database.* J Am Soc Mass Spectrom, 1994. **5**(11): p. 976-89.
151. Perkins, D.N., D.J. Pappin, D.M. Creasy, and J.S. Cottrell, *Probability-based protein identification by searching sequence databases using mass spectrometry data.* Electrophoresis, 1999. **20**(18): p. 3551-67.
152. Dorfer, V., P. Pichler, T. Stranzl, J. Stadlmann, T. Taus, S. Winkler, and K. Mechtler, *MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra.* J Proteome Res, 2014. **13**(8): p. 3679-84.
153. Craig, R. and R.C. Beavis, *TANDEM: matching proteins with tandem mass spectra.* Bioinformatics, 2004. **20**(9): p. 1466-7.
154. Cox, J., N. Neuhauser, A. Michalski, R.A. Scheltema, J.V. Olsen, and M. Mann, *Andromeda: a peptide search engine integrated into the MaxQuant environment.* J Proteome Res, 2011. **10**(4): p. 1794-805.
155. Kim, S., N. Mischerikow, N. Bandeira, J.D. Navarro, L. Wich, S. Mohammed, A.J. Heck, and P.A. Pevzner, *The generating function of CID, ETD, and CID/ETD pairs of tandem mass spectra: applications to database search.* Mol Cell Proteomics, 2010. **9**(12): p. 2840-52.
156. Bern, M., Y.J. Kil, and C. Becker, *Byonic: advanced peptide and protein identification software.* Curr Protoc Bioinformatics, 2012. **Chapter 13**: p. Unit13 20.
157. Kall, L., J.D. Canterbury, J. Weston, W.S. Noble, and M.J. MacCoss, *Semi-supervised learning for peptide identification from shotgun proteomics datasets.* Nat Methods, 2007. **4**(11): p. 923-5.
158. Elias, J.E. and S.P. Gygi, *Target-decoy search strategy for mass spectrometry-based proteomics.* Methods Mol Biol, 2010. **604**: p. 55-71.

159. Paik, Y.K., S.K. Jeong, G.S. Omenn, M. Uhlen, S. Hanash, S.Y. Cho, H.J. Lee, K. Na, E.Y. Choi, F. Yan, F. Zhang, Y. Zhang, M. Snyder, Y. Cheng, R. Chen, G. Marko-Varga, E.W. Deutsch, H. Kim, J.Y. Kwon, R. Aebersold, A. Bairoch, A.D. Taylor, K.Y. Kim, E.Y. Lee, D. Hochstrasser, P. Legrain, and W.S. Hancock, *The Chromosome-Centric Human Proteome Project for cataloging proteins encoded in the genome.* Nat Biotechnol, 2012. **30**(3): p. 221-3.
160. Gaudet, P., G. Argoud-Puy, I. Cusin, P. Duek, O. Evalet, A. Gateau, A. Gleizes, M. Pereira, M. Zahn-Zabal, C. Zwahlen, A. Bairoch, and L. Lane, *neXtProt: organizing protein knowledge in the context of human proteome projects.* J Proteome Res, 2013. **12**(1): p. 293-8.
161. Geer, L.Y., A. Marchler-Bauer, R.C. Geer, L. Han, J. He, S. He, C. Liu, W. Shi, and S.H. Bryant, *The NCBI BioSystems database.* Nucleic Acids Research, 2010. **38**(suppl 1): p. D492-D496.
162. Craig, R., J.P. Cortens, and R.C. Beavis, *Open source system for analyzing, validating, and storing protein identification data.* J Proteome Res, 2004. **3**(6): p. 1234-42.
163. Wang, M., C.J. Herrmann, M. Simonovic, D. Szklarczyk, and C. von Mering, *Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines.* Proteomics, 2015.
164. Farrah, T., E.W. Deutsch, G.S. Omenn, Z. Sun, J.D. Watts, T. Yamamoto, D. Shteynberg, M.M. Harris, and R.L. Moritz, *State of the human proteome in 2013 as viewed through PeptideAtlas: comparing the kidney, urine, and plasma proteomes for the biology- and disease-driven Human Proteome Project.* J Proteome Res, 2014. **13**(1): p. 60-75.
165. Vizcaino, J.A., R.G. Cote, A. Csordas, J.A. Dianes, A. Fabregat, J.M. Foster, J. Griss, E. Alpi, M. Birim, J. Contell, G. O'Kelly, A. Schoenegger, D. Ovelleiro, Y. Perez-Riverol, F. Reisinger, D. Rios, R. Wang, and H. Hermjakob, *The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013.* Nucleic Acids Res, 2013. **41**(Database issue): p. D1063-9.
166. Vizcaino, J.A., E.W. Deutsch, R. Wang, A. Csordas, F. Reisinger, D. Rios, J.A. Dianes, Z. Sun, T. Farrah, N. Bandeira, P.A. Binz, I. Xenarios, M. Eisenacher, G. Mayer, L. Gatto, A. Campos, R.J. Chalkley, H.J. Kraus, J.P. Albar, S. Martinez-Bartolome, R. Apweiler, G.S. Omenn, L. Martens, A.R. Jones, and H. Hermjakob, *ProteomeXchange provides globally coordinated proteomics data submission and dissemination.* Nat Biotechnol, 2014. **32**(3): p. 223-6.
167. Ong, S.E. and M. Mann, *Mass spectrometry-based proteomics turns quantitative.* Nat Chem Biol, 2005. **1**(5): p. 252-62.
168. Bantscheff, M., S. Lemeer, M.M. Savitski, and B. Kuster, *Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present.* Anal Bioanal Chem, 2012. **404**(4): p. 939-65.
169. Chahrour, O., D. Cobice, and J. Malone, *Stable isotope labelling methods in mass spectrometry-based quantitative proteomics.* J Pharm Biomed Anal, 2015.
170. Megger, D.A., T. Bracht, H.E. Meyer, and B. Sitek, *Label-free quantification in clinical proteomics.* Biochim Biophys Acta, 2013. **1834**(8): p. 1581-90.

171. Picotti, P. and R. Aebersold, *Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions.* Nat Methods, 2012. **9**(6): p. 555-66.
172. Wasinger, V.C., M. Zeng, and Y. Yau, *Current status and advances in quantitative proteomic mass spectrometry.* Int J Proteomics, 2013. **2013**: p. 180605.
173. Zhu, W., J.W. Smith, and C.M. Huang, *Mass spectrometry-based label-free quantitative proteomics.* J Biomed Biotechnol, 2010. **2010**: p. 840518.
174. Liu, H., R.G. Sadygov, and J.R. Yates, 3rd, *A model for random sampling and estimation of relative protein abundance in shotgun proteomics.* Analytical Chemistry, 2004. **76**(14): p. 4193-201.
175. Bondarenko, P.V., D. Chelius, and T.A. Shaler, *Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry.* Anal Chem, 2002. **74**(18): p. 4741-9.
176. Chelius, D. and P.V. Bondarenko, *Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry.* J Proteome Res, 2002. **1**(4): p. 317-23.
177. Li, Z., R.M. Adams, K. Chourey, G.B. Hurst, R.L. Hettich, and C. Pan, *Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ Orbitrap Velos.* J Proteome Res, 2012. **11**(3): p. 1582-90.
178. Merl, J., M. Ueffing, S.M. Hauck, and C. von Toerne, *Direct comparison of MS-based label-free and SILAC quantitative proteome profiling strategies in primary retinal Muller cells.* Proteomics, 2012. **12**(12): p. 1902-11.
179. Sjodin, M.O., M. Wetterhall, K. Kultima, and K. Artemenko, *Comparative study of label and label-free techniques using shotgun proteomics for relative protein quantification.* J Chromatogr B Analyt Technol Biomed Life Sci, 2013. **928**: p. 83-92.
180. Zhang, R. and F.E. Regnier, *Minimizing resolution of isotopically coded peptides in comparative proteomics.* J Proteome Res, 2002. **1**(2): p. 139-47.
181. Ong, S.E., B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, and M. Mann, *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.* Mol Cell Proteomics, 2002. **1**(5): p. 376-86.
182. Chen, X., S. Wei, Y. Ji, X. Guo, and F. Yang, *Quantitative proteomics using SILAC: principles, applications and developments.* Proteomics, 2015.
183. Gygi, S.P., B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, and R. Aebersold, *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags.* Nat Biotechnol, 1999. **17**(10): p. 994-9.
184. Hsu, J.L., S.Y. Huang, N.H. Chow, and S.H. Chen, *Stable-isotope dimethyl labeling for quantitative proteomics.* Anal Chem, 2003. **75**(24): p. 6843-52.
185. Boersema, P.J., R. Raijmakers, S. Lemeer, S. Mohammed, and A.J. Heck, *Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics.* Nat Protoc, 2009. **4**(4): p. 484-94.
186. Ross, P.L., Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, and D.J. Pappin, *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-*

reactive isobaric tagging reagents. Mol Cell Proteomics, 2004. **3**(12): p. 1154-69.
187. Thompson, A., J. Schafer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T. Neumann, R. Johnstone, A.K. Mohammed, and C. Hamon, *Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS.* Anal Chem, 2003. **75**(8): p. 1895-904.
188. Choe, L., M. D'Ascenzo, N.R. Relkin, D. Pappin, P. Ross, B. Williamson, S. Guertin, P. Pribil, and K.H. Lee, *8-plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease.* Proteomics, 2007. **7**(20): p. 3651-60.
189. Werner, T., G. Sweetman, M.F. Savitski, T. Mathieson, M. Bantscheff, and M.M. Savitski, *Ion coalescence of neutron encoded TMT 10-plex reporter ions.* Anal Chem, 2014. **86**(7): p. 3594-601.
190. Villanueva, J., M. Carrascal, and J. Abian, *Isotope dilution mass spectrometry for absolute quantification in proteomics: concepts and strategies.* J Proteomics, 2014. **96**: p. 184-99.
191. Gerber, S.A., J. Rush, O. Stemman, M.W. Kirschner, and S.P. Gygi, *Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS.* Proc Natl Acad Sci U S A, 2003. **100**(12): p. 6940-5.
192. Beynon, R.J., M.K. Doherty, J.M. Pratt, and S.J. Gaskell, *Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides.* Nat Methods, 2005. **2**(8): p. 587-9.
193. Brun, V., A. Dupuis, A. Adrait, M. Marcellin, D. Thomas, M. Court, F. Vandenesch, and J. Garin, *Isotope-labeled protein standards: toward absolute quantitative proteomics.* Mol Cell Proteomics, 2007. **6**(12): p. 2139-49.
194. Hanke, S., H. Besir, D. Oesterhelt, and M. Mann, *Absolute SILAC for accurate quantitation of proteins in complex mixtures down to the attomole level.* J Proteome Res, 2008. **7**(3): p. 1118-30.
195. Singh, S., M. Springer, J. Steen, M.W. Kirschner, and H. Steen, *FLEXIQuant: a novel tool for the absolute quantification of proteins, and the simultaneous identification and quantification of potentially modified peptides.* J Proteome Res, 2009. **8**(5): p. 2201-10.
196. Zinn, N., D. Winter, and W.D. Lehmann, *Recombinant isotope labeled and selenium quantified proteins for absolute protein quantification.* Anal Chem, 2010. **82**(6): p. 2334-40.
197. Zeiler, M., W.L. Straube, E. Lundberg, M. Uhlen, and M. Mann, *A Protein Epitope Signature Tag (PrEST) library allows SILAC-based absolute quantification and multiplexed determination of protein copy numbers in cell lines.* Mol Cell Proteomics, 2012. **11**(3): p. O111 009613.
198. Zeiler, M., M. Moser, and M. Mann, *Copy number analysis of the murine platelet proteome spanning the complete abundance range.* Mol Cell Proteomics, 2014. **13**(12): p. 3435-45.
199. Picotti, P., H. Lam, D. Campbell, E. Deutsch, H. Mirzaei, J. Ranish, B. Domon, and R. Aebersold, *A database of validated assays for the targeted mass spectrometric analysis of the S. cerevisiae proteome.* Nat Methods, 2008. **5**(11): p. 913-4.

200. Mead, J.A., L. Bianco, V. Ottone, C. Barton, R.G. Kay, K.S. Lilley, N.J. Bond, and C. Bessant, *MRMaid, the web-based tool for designing multiple reaction monitoring (MRM) transitions.* Mol Cell Proteomics, 2009. **8**(4): p. 696-705.
201. Martin, D.B., T. Holzman, D. May, A. Peterson, A. Eastham, J. Eng, and M. McIntosh, *MRMer, an interactive open source and cross-platform system for data extraction and visualization of multiple reaction monitoring experiments.* Mol Cell Proteomics, 2008. **7**(11): p. 2270-8.
202. Sherwood, C.A., A. Eastham, L.W. Lee, A. Peterson, J.K. Eng, D. Shteynberg, L. Mendoza, E.W. Deutsch, J. Risler, N. Tasman, R. Aebersold, H. Lam, and D.B. Martin, *MaRiMba: a software application for spectral library-based MRM transition list assembly.* J Proteome Res, 2009. **8**(10): p. 4396-405.
203. Gallien, S. and B. Domon, *Detection and quantification of proteins in clinical samples using high resolution mass spectrometry.* Methods, 2015. **81**: p. 15-23.
204. Gallien, S., E. Duriez, C. Crone, M. Kellmann, T. Moehring, and B. Domon, *Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer.* Mol Cell Proteomics, 2012. **11**(12): p. 1709-23.
205. Peterson, A.C., J.D. Russell, D.J. Bailey, M.S. Westphall, and J.J. Coon, *Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics.* Mol Cell Proteomics, 2012. **11**(11): p. 1475-88.
206. MacLean, B., D.M. Tomazela, N. Shulman, M. Chambers, G.L. Finney, B. Frewen, R. Kern, D.L. Tabb, D.C. Liebler, and M.J. MacCoss, *Skyline: an open source document editor for creating and analyzing targeted proteomics experiments.* Bioinformatics, 2010. **26**(7): p. 966-8.
207. Gallien, S., S.Y. Kim, and B. Domon, *Large-Scale Targeted Proteomics Using Internal Standard Triggered-Parallel Reaction Monitoring (IS-PRM).* Mol Cell Proteomics, 2015. **14**(6): p. 1630-44.
208. Kuzyk, M.A., D. Smith, J. Yang, T.J. Cross, A.M. Jackson, D.B. Hardie, N.L. Anderson, and C.H. Borchers, *Multiple reaction monitoring-based, multiplexed, absolute quantitation of 45 proteins in human plasma.* Mol Cell Proteomics, 2009. **8**(8): p. 1860-77.
209. Shi, T., X. Sun, Y. Gao, T.L. Fillmore, A.A. Schepmoes, R. Zhao, J. He, R.J. Moore, J. Kagan, K.D. Rodland, T. Liu, A.Y. Liu, R.D. Smith, K. Tang, D.G. Camp, 2nd, and W.J. Qian, *Targeted quantification of low ng/mL level proteins in human serum without immunoaffinity depletion.* J Proteome Res, 2013. **12**(7): p. 3353-61.
210. Anderson, L., *Six decades searching for meaning in the proteome.* J Proteomics, 2014. **107**: p. 24-30.
211. Rifai, N., M.A. Gillette, and S.A. Carr, *Protein biomarker discovery and validation: the long and uncertain path to clinical utility.* Nat Biotechnol, 2006. **24**(8): p. 971-83.
212. Surinova, S., R. Schiess, R. Huttenhain, F. Cerciello, B. Wollscheid, and R. Aebersold, *On the development of plasma protein biomarkers.* J Proteome Res, 2011. **10**(1): p. 5-16.
213. Arentz, G., F. Weiland, M.K. Oehler, and P. Hoffmann, *State of the art of 2D DIGE.* Proteomics Clin Appl, 2015. **9**(3-4): p. 277-88.

214. Dowell, J.A., D.C. Frost, J. Zhang, and L. Li, *Comparison of two-dimensional fractionation techniques for shotgun proteomics.* Anal Chem, 2008. **80**(17): p. 6715-23.
215. Bonifacino, J.S., E.C. Dell'Angelica, and T.A. Springer, *Immunoprecipitation.* Curr Protoc Immunol, 2001. **Chapter 8**: p. Unit 8 3.
216. Nelson, R.W., J.R. Krone, A.L. Bieber, and P. Williams, *Mass Spectrometric Immunoassay.* Analytical Chemistry, 1995. **67**(7): p. 1153-1158.
217. Anderson, N.L., N.G. Anderson, L.R. Haines, D.B. Hardie, R.W. Olafson, and T.W. Pearson, *Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA).* J Proteome Res, 2004. **3**(2): p. 235-44.
218. Weiss, F., B.H. van den Berg, H. Planatscher, C.J. Pynn, T.O. Joos, and O. Poetz, *Catch and measure-mass spectrometry-based immunoassays in biomarker research.* Biochim Biophys Acta, 2014. **1844**(5): p. 927-32.
219. Ekstrom, S., L. Wallman, J. Malm, C. Becker, H. Lilja, T. Laurell, and G. Marko-Varga, *Integrated selective enrichment target--a microtechnology platform for matrix-assisted laser desorption/ionization-mass spectrometry applied on protein biomarkers in prostate diseases.* Electrophoresis, 2004. **25**(21-22): p. 3769-77.
220. Ahmad-Tajudin, A., B. Adler, S. Ekstrom, G. Marko-Varga, J. Malm, H. Lilja, and T. Laurell, *MALDI-target integrated platform for affinity-captured protein digestion.* Anal Chim Acta, 2014. **807**: p. 1-8.
221. Thulasiraman, V., S. Lin, L. Gheorghiu, J. Lathrop, L. Lomas, D. Hammond, and E. Boschetti, *Reduction of the concentration difference of proteins in biological liquids using a library of combinatorial ligands.* Electrophoresis, 2005. **26**(18): p. 3561-3571.
222. Wingren, C., P. James, and C.A.K. Borrebaeck, *Strategy for surveying the proteome using affinity proteomics and mass spectrometry.* Proteomics, 2009. **9**(6): p. 1511-1517.
223. Olsson, N., C. Wingren, M. Mattsson, P. James, D. O'Connell, F. Nilsson, D.J. Cahill, and C.A. Borrebaeck, *Proteomic analysis and discovery using affinity proteomics and mass spectrometry.* Mol Cell Proteomics, 2011. **10**(10): p. M110 003962.
224. Poetz, O., S. Hoeppe, M.F. Templin, D. Stoll, and T.O. Joos, *Proteome wide screening using peptide affinity capture.* Proteomics, 2009. **9**(6): p. 1518-1523.
225. Hoeppe, S., T.D. Schreiber, H. Planatscher, A. Zell, M.F. Templin, D. Stoll, T.O. Joos, and O. Poetz, *Targeting peptide termini, a novel immunoaffinity approach to reduce complexity in mass spectrometric protein identification.* Mol Cell Proteomics, 2011. **10**(2): p. M110 002857.
226. Fang, X., L. Huang, J.S. Feitelson, and W.W. Zhang, *Affinity separation: divide and conquer the proteome.* Drug Discov Today Technol, 2004. **1**(2): p. 141-8.
227. Fang, X. and W.W. Zhang, *Affinity separation and enrichment methods in proteomic analysis.* J Proteomics, 2008. **71**(3): p. 284-303.
228. Ly, L. and V.C. Wasinger, *Protein and peptide fractionation, enrichment and depletion: tools for the complex proteome.* Proteomics, 2011. **11**(4): p. 513-34.

229. Selvaraju, S. and Z.E. Rassi, *Liquid-phase-based separation systems for depletion, prefractionation and enrichment of proteins in biological fluids and matrices for in-depth proteomics analysis--an update covering the period 2008-2011.* Electrophoresis, 2012. **33**(1): p. 74-88.
230. Zhang, W.W., *The use of gene-specific IgY antibodies for drug target discovery.* Drug Discov Today, 2003. **8**(8): p. 364-71.
231. Caironi, P. and L. Gattinoni, *The clinical use of albumin: the point of view of a specialist in intensive care.* Blood Transfus, 2009. **7**(4): p. 259-67.
232. Bellei, E., S. Bergamini, E. Monari, L.I. Fantoni, A. Cuoghi, T. Ozben, and A. Tomasi, *High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins.* Amino Acids, 2011. **40**(1): p. 145-56.
233. Granger, J., J. Siddiqui, S. Copeland, and D. Remick, *Albumin depletion of human plasma also removes low abundance proteins including the cytokines.* Proteomics, 2005. **5**(18): p. 4713-8.
234. Roche, S., L. Tiers, M. Provansal, M. Seveno, M.T. Piva, P. Jouin, and S. Lehmann, *Depletion of one, six, twelve or twenty major blood proteins before proteomic analysis: the more the better?* J Proteomics, 2009. **72**(6): p. 945-51.
235. Guldbrandsen, A., H. Vethe, Y. Farag, E. Oveland, H. Garberg, M. Berle, K.M. Myhr, J.A. Opsahl, H. Barsnes, and F.S. Berven, *In-depth Characterization of the Cerebrospinal Fluid (CSF) Proteome Displayed Through the CSF Proteome Resource (CSF-PR).* Mol Cell Proteomics, 2014. **13**(11): p. 3152-63.
236. Schutzer, S.E., T. Liu, B.H. Natelson, T.E. Angel, A.A. Schepmoes, S.O. Purvine, K.K. Hixson, M.S. Lipton, D.G. Camp, P.K. Coyle, R.D. Smith, and J. Bergquist, *Establishing the proteome of normal human cerebrospinal fluid.* PLoS One, 2010. **5**(6): p. e10980.
237. Gunther, R., E. Krause, M. Schumann, J. Ausseil, J.M. Heard, I.E. Blasig, and R.F. Haseloff, *Removal of albumin and immunoglobulins from canine cerebrospinal fluid using depletion kits: a feasibility study.* Fluids Barriers CNS, 2014. **11**: p. 14.
238. Hinerfeld, D., D. Innamorati, J. Pirro, and S.W. Tam, *Serum/Plasma depletion with chicken immunoglobulin Y antibodies for proteomic analysis from multiple Mammalian species.* J Biomol Tech, 2004. **15**(3): p. 184-90.
239. Cunningham, R., P. Jany, A. Messing, and L. Li, *Protein changes in immunodepleted cerebrospinal fluid from a transgenic mouse model of Alexander disease detected using mass spectrometry.* J Proteome Res, 2013. **12**(2): p. 719-28.
240. Linke, T., S. Doraiswamy, and E.H. Harrison, *Rat plasma proteomics: effects of abundant protein depletion on proteomic analysis.* J Chromatogr B Analyt Technol Biomed Life Sci, 2007. **849**(1-2): p. 273-81.
241. Jeong, J.S., L. Jiang, E. Albino, J. Marrero, H.S. Rho, J. Hu, S. Hu, C. Vera, D. Bayron-Poueymiroy, Z.A. Rivera-Pacheco, L. Ramos, C. Torres-Castro, J. Qian, J. Bonaventura, J.D. Boeke, W.Y. Yap, I. Pino, D.J. Eichinger, H. Zhu, and S. Blackshaw, *Rapid identification of monospecific monoclonal antibodies using a human proteome microarray.* Mol Cell Proteomics, 2012. **11**(6): p. O111 016253.

242. Jia, X.C., R. Raya, L. Zhang, O. Foord, W.L. Walker, M.L. Gallo, M. Haak-Frendscho, L.L. Green, and C.G. Davis, *A novel method of Multiplexed Competitive Antibody Binning for the characterization of monoclonal antibodies.* J Immunol Methods, 2004. **288**(1-2): p. 91-8.
243. Michaud, G.A., M. Salcius, F. Zhou, R. Bangham, J. Bonin, H. Guo, M. Snyder, P.F. Predki, and B.I. Schweitzer, *Analyzing antibody specificity with whole proteome microarrays.* Nat Biotechnol, 2003. **21**(12): p. 1509-12.
244. Poetz, O., R. Ostendorp, B. Brocks, J.M. Schwenk, D. Stoll, T.O. Joos, and M.F. Templin, *Protein microarrays for antibody profiling: specificity and affinity determination on a chip.* Proteomics, 2005. **5**(9): p. 2402-11.
245. Predki, P.F., D. Mattoon, R. Bangham, B. Schweitzer, and G. Michaud, *Protein microarrays: a new tool for profiling antibody cross-reactivity.* Hum Antibodies, 2005. **14**(1-2): p. 7-15.
246. Machida, K., C.M. Thompson, K. Dierck, K. Jablonowski, S. Karkkainen, B. Liu, H. Zhang, P.D. Nash, D.K. Newman, P. Nollau, T. Pawson, G.H. Renkema, K. Saksela, M.R. Schiller, D.G. Shin, and B.J. Mayer, *High-throughput phosphotyrosine profiling using SH2 domains.* Mol Cell, 2007. **26**(6): p. 899-915.
247. Scott, J.D. and T. Pawson, *Cell signaling in space and time: where proteins come together and when they're apart.* Science, 2009. **326**(5957): p. 1220-4.
248. Seet, B.T., I. Dikic, M.M. Zhou, and T. Pawson, *Reading protein modifications with interaction domains.* Nat Rev Mol Cell Biol, 2006. **7**(7): p. 473-83.
249. Structural Genomics, C., C. China Structural Genomics, C. Northeast Structural Genomics, S. Graslund, P. Nordlund, J. Weigelt, B.M. Hallberg, J. Bray, O. Gileadi, S. Knapp, U. Oppermann, C. Arrowsmith, R. Hui, J. Ming, S. dhe-Paganon, H.W. Park, A. Savchenko, A. Yee, A. Edwards, R. Vincentelli, C. Cambillau, R. Kim, S.H. Kim, Z. Rao, Y. Shi, T.C. Terwilliger, C.Y. Kim, L.W. Hung, G.S. Waldo, Y. Peleg, S. Albeck, T. Unger, O. Dym, J. Prilusky, J.L. Sussman, R.C. Stevens, S.A. Lesley, I.A. Wilson, A. Joachimiak, F. Collart, I. Dementieva, M.I. Donnelly, W.H. Eschenfeldt, Y. Kim, L. Stols, R. Wu, M. Zhou, S.K. Burley, J.S. Emtage, J.M. Sauder, D. Thompson, K. Bain, J. Luz, T. Gheyi, F. Zhang, S. Atwell, S.C. Almo, J.B. Bonanno, A. Fiser, S. Swaminathan, F.W. Studier, M.R. Chance, A. Sali, T.B. Acton, R. Xiao, L. Zhao, L.C. Ma, J.F. Hunt, L. Tong, K. Cunningham, M. Inouye, S. Anderson, H. Janjua, R. Shastry, C.K. Ho, D. Wang, H. Wang, M. Jiang, G.T. Montelione, D.I. Stuart, R.J. Owens, S. Daenke, A. Schutz, U. Heinemann, S. Yokoyama, K. Bussow, and K.C. Gunsalus, *Protein production and purification.* Nat Methods, 2008. **5**(2): p. 135-46.
250. Woodruff-Pak, D.S., *Animal models of Alzheimer's disease: therapeutic implications.* J Alzheimers Dis, 2008. **15**(4): p. 507-21.
251. Packer, R.M., N.K. Shihab, B.B. Torres, and H.A. Volk, *Clinical risk factors associated with anti-epileptic drug responsiveness in canine epilepsy.* PLoS One, 2014. **9**(8): p. e106026.
252. Head, E., *A canine model of human aging and Alzheimer's disease.* Biochim Biophys Acta, 2013. **1832**(9): p. 1384-9.

253. Gianazza, E., E. Vegeto, I. Eberini, C. Sensi, and I. Miller, *Neglected markers: altered serum proteome in murine models of disease.* Proteomics, 2012. **12**(4-5): p. 691-707.
254. Gianazza, E., R. Wait, I. Eberini, C. Sensi, L. Sironi, and I. Miller, *Proteomics of rat biological fluids--the tenth anniversary update.* J Proteomics, 2012. **75**(11): p. 3113-28.
255. Lieschke, G.J. and P.D. Currie, *Animal models of human disease: zebrafish swim into view.* Nat Rev Genet, 2007. **8**(5): p. 353-67.
256. Bendixen, E., M. Danielsen, K. Hollung, E. Gianazza, and I. Miller, *Farm animal proteomics--a review.* J Proteomics, 2011. **74**(3): p. 282-93.
257. Eckersall, P.D., A.M. de Almeida, and I. Miller, *Proteomics, a new tool for farm animal science.* J Proteomics, 2012. **75**(14): p. 4187-9.
258. Soares, R., C. Franco, E. Pires, M. Ventosa, R. Palhinhas, K. Koci, A. Martinho de Almeida, and A. Varela Coelho, *Mass spectrometry and animal science: protein identification strategies and particularities of farm animal species.* J Proteomics, 2012. **75**(14): p. 4190-206.
259. Neiman, M., C. Fredolini, H. Johansson, J. Lehtio, P.A. Nygren, M. Uhlen, P. Nilsson, and J.M. Schwenk, *Selectivity analysis of single binder assays used in plasma protein profiling.* Proteomics, 2013. **13**(23-24): p. 3406-10.
260. Qundos, U., H. Johannesson, C. Fredolini, G. O'Hurley, R. Branca, M. Uhlén, F. Wiklund, A. Bjartell, P. Nilsson, and J.M. Schwenk, *Analysis of plasma from prostate cancer patients links decreased carnosine dipeptidase 1 levels to lymph node metastasis.* Advances in Integrative Medicine, 2014. **2**(0): p. 14-24.
261. Prakash, A., T. Rezai, B. Krastins, D. Sarracino, M. Athanas, P. Russo, M.M. Ross, H. Zhang, Y. Tian, V. Kulasingam, A.P. Drabovich, C. Smith, I. Batruch, L. Liotta, E. Petricoin, E.P. Diamandis, D.W. Chan, and M.F. Lopez, *Platform for establishing interlaboratory reproducibility of selected reaction monitoring-based mass spectrometry peptide assays.* J Proteome Res, 2010. **9**(12): p. 6678-88.
262. Scherl, A., *Clinical protein mass spectrometry.* Methods, 2015. **81**: p. 3-14.
263. Strathmann, F.G. and A.N. Hoofnagle, *Current and future applications of mass spectrometry to the clinical laboratory.* Am J Clin Pathol, 2011. **136**(4): p. 609-16.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1272

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-259623

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2015